华章数学译丛

19

Wavelet Methods for Time Series Analysis

时间序列分析的小波方法

(美) Donald B. Percival
(英) Andrew T. Walden 著
程正兴 等译



机械工业出版社
China Machine Press

时间序列分析的小波方法

时间序列分析是用随机过程理论和数理统计学的方法，研究随机数据序列所遵从的统计规律，用于解决科研、工程技术、金融及经济等诸多领域内的实际问题。

本书是一本由浅入深的小波分析导论，介绍了基于小波的时间序列统计分析。实践中的离散时间技术是本书的论述重点，同时对于理解和实现离散小波变换将涉及的诸多原理与算法也进行了详细的描述。

本书网站（http://www.staff.washington.edu/dbp/wmtsa.html）有所用时间序列与小波的材料，并可以得到用S-Plus和其他语言开发软件的信息等。

作者简介



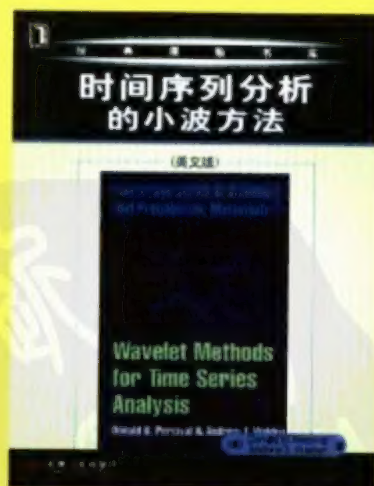
Donald B. Percival 于1983年在华盛顿大学获得统计学博士学位。目前是华盛顿大学应用物理实验室首席数学家，并且是华盛顿大学统计系副教授。他的研究方向包括谱分析、自然科学中的小波和统计方法的应用。除本书外，他还著有《Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques》一书。他还是《Journal of Computational and Graphical Statistics》（该杂志由美国统计学会主办）的编辑。



Andrew T. Walden 伦敦大学帝国科技医学学院数学系教授。其研究方向为非平稳时间序列中的离散和连续小波变换的理论及应用，时间序列（特别是复值时间序列）在地球物理学、海洋学和医学中的应用等。除本书外，他还写有很多论文。

Wavelet Methods for Time Series Analysis





影印版
ISBN 7-111-14118-0
定价：58.00 元





华章图书

华章网站 http://www.hzbook.com

网上购书：www.china-pub.com

投稿热线：(010) 88379604
购书热线：(010) 68995259, 68995264
读者信箱：hzjsj@hzbook.com

ISBN 7-111-17806-8/O・464
定价：69.00 元

封面设计・杨宇梅


自助
書屋

华章数学译丛

19

Wavelet Methods for Time Series Analysis

时间序列分析的小波方法

(美) Donald B. Percival
(英) Andrew T. Walden 著
程正兴 等译



机械工业出版社
China Machine Press

本书详细地介绍了小波方法在时间序列分析中的应用，图例丰富，语言简明易懂，论述严谨. 另外，本书对小波分析所需要的数学知识进行了简洁实用的讲解，还在正文中嵌入了大量的练习，并在附录中给出了这些练习的答案，同时每章另备有适于课堂布置的练习.

本书适合作为高等院校统计学、数学等专业学生的教材，同时也可作为从事相关领域研究的人员的参考书.

Donald B. Percival and Andrew T. Walden: Wavelet Methods for Time Series Analysis (ISBN 0-521-64068-7).

Originally published by Cambridge University Press in 2000.

This Chinese edition is published with the permission of the Syndicate of the Press of the University of Cambridge, Cambridge, England.

本书版权登记号：图字：01-2004-1664

图书在版编目(CIP)数据

时间序列分析的小波方法/(美)珀西瓦尔(Percival, D. B.)等著；程正兴等译. —北京：机械工业出版社，2006.3

(华章数学译丛)

书名原文：Wavelet Methods for Time Series Analysis

ISBN 7-111-17806-8

Ⅰ. 时…　Ⅱ. ①珀…　②程…　Ⅲ. 小波分析—应用—时间序列分析　Ⅳ. O211.61

中国版本图书馆 CIP 数据核字(2005)第 128312 号

机械工业出版社(北京市西城区百万庄大街 22 号　邮政编码　100037)

责任编辑：白红莉　迟振春

北京牛山世兴印刷厂印刷 · 新华书店北京发行所发行

2006 年 3 月第 1 版第 1 次印刷

787mm×1020mm 1/16 · 36.25 印张

印数：0 001-4 000 册

定价：69.00 元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

本社购书热线：(010)68326294

译者序

2001 年底，我在小波网上看到过这本书的简介，但始终没有见到它，2004 年，机械工业出版社邀请我翻译这本书，由于有久别重逢的感觉，所以毅然接受.

本书是剑桥大学出版社于 2000 年出版的为研究生课程写的一本教材，书中有许多内容是作者长期研究的结果. 本书的最大特点就是把一元离散小波变换进行演绎，在离散时间上与应用十分广泛的时间序列分析有机地结合在一起，这是小波分析应用的一种创造性尝试，可以看作是小波应用于实际的一个典范.

自己与本书作者从未谋面，但本书的写作手法无疑会给从事小波分析应用研究的人们一种全新的感觉. 书中把小波分析中的离散小波变换和离散小波包变换以及它们的变形应用于时间序列的分析，定位准确、分析细致、论述充分、说服力强，尤其值得肯定的是本书对实际例子的分析，真正在理论和实践之间架起了一座桥梁.

本书还是自我包容的，对应用所需的离散小波变换、离散小波包变换、多分辨分析、统计与随机过程简介等都有详细的讲述，另外书中还配备了帮助进一步理解的嵌入练习以及章后练习.

本书可作为理工科研究生(包括硕士生和博士生)学习小波分析与应用的教材，也适合希望研究小波分析应用的研究者，特别适合统计、电子工程、物理学、地球物理学、天文学、海洋学以及其他物理学科的研究者. 对于小波分析理论与算法的研究者来说，研读本书必会对以后的研究工作带来许多好处.

本书由于篇幅大，所以翻译时由王郑耀(第 1、7 章)、仝黎(第 2、3 章)、郭潇湧(第 4 章)、周雪芹(第 5、6 章)、郭丽娜(第 8 章)、徐亮(第 9 章)、孙传姣(第 10 章)、吴梦秋(第 11 章)、谭小龙(附录)对各章先进行初译，然后由我反复改过三次. 尤其是谭小龙，除初译外，还做了许多修改打印等方面的工作，在这里向他们表示感谢. 本书我们采用直译，以保持原书的特有风格. 虽然本书的翻译花费了大量心血，但还是会出现错误，希望广大读者不吝指正.

程正兴

于西安交通大学理学院

2005 年 5 月

前　言

近十年来人们对小波的兴趣急剧增加，涉及的领域包括数学、物理学、电子工程以及其他学科．结果是，小波方法在诸如微分方程、图像处理和统计等不同的领域中具有重要的影响．本书是离散时间序列分析中有关小波分析及其应用的一个导引，这些时间序列是自然科学中需要的典型序列．而我们全面介绍作为离散小波变换(DWT)后盾的基本理论，目的是在理论与实践之间架起一座桥梁．本书的主要内容包括：

- 以实践的术语强调离散小波变换的实际含义．
- 展示离散小波变换如何用于产生时间序列分析学家需要的信息描述统计．
- 讨论随机模型如何用于评估由离散小波变换计算的量的统计性质．
- 提供自然科学中有代表性的时间序列的小波分析的大量例子．

到现在为止，关于小波的大多数书籍都是按照连续函数来展开描述的，并且常常介绍给读者过多的不同类型的小波．而本书则通过标准的滤波和矩阵变换的思想，在离散时间上讲述小波方法．为避免读者负担过重，本书总是集中于Daubechies(1992)描述的小波滤波器类上，这个小波滤波器类对于统计的应用特别方便和有用；然而，通过研究 Daubechies 小波类，读者可以很好地理解其他感兴趣的小波类．为了教学的目的，事实上本书在开头(第 1 章)和结束(第 11 章)都讨论连续的情形．这种方式允许我们一开始从历史视角激发好的思想，然后在最后把离散分析中的思想结合到连续时间小波分析中的一些已知结果中．

在本书中较早叙述的论题(第 4、5 章)包括离散小波变换(DWT)和极大重叠离散小波变换(MODWT)，而极大重叠离散小波变换可以认为是具有吸引人的性质的离散小波变换的推广．总的来看，这两章强调了计算离散小波变换和极大重叠离散小波变换的算法，以及这些变换如何用于提供时间序列的信息描述统计．特别地，两个变换都导致了时间序列的抽样方差基于尺度的分解，以及基于尺度的加性分解(称之为多分辨分析)．离散小波变换和极大重叠离散小波变换的推广(即文献中所称的“小波包”变换)和时间序列经由匹配追踪的分解，是第 6 章的主题．在本书的第二部分，我们把这些变换与随机模型结合起来，以提出时间序列分析的基于小波的统计推断．书中这部分涉及的特定主题包括：

- 小波方差，它提供了一个基于尺度的方差分析，与传统的基于频率的谱分析互补(第 8 章)．
- “长记忆过程”(即具有慢速衰减相关的过程)的分析和综合(第 9 章)．
- 经由“阈值”和“去噪”的信号估计(第 10 章)．

本书的写作方式是“由基础一层一层上升”(由浅入深)．我们试图使本书尽可能是自我包容的(为此，第 2、3、7 章分别包括回顾傅里叶理论和滤波理论；时

间序列的规范正交变换中的关键思想；涉及随机变量和随机过程的重要概念）. 因此，本书适用于优秀的大学本科生，但主要是为研究生以及统计学、电子工程、物理学、地球物理学、天文学、海洋学等其他自然科学的研究者写的. 具有较强数学基础的读者可以快速阅读第 2、3 章. 具有离散小波变换知识的读者可以利用第 4、5 章各节后的关键结论和定义来确定需要掌握的内容. 本书的草稿在华盛顿大学用作研究生课程的教科书已经五年了，但是我们还是设计它是自学参考书而包含嵌入到各章中的大量练习(特别是第 2 章到第 5 章)，而在附录中提供了解答. 嵌入练习的目的是方便读者逐步理解书中内容. 为了作为教科书使用，我们在各章的最后还提供了附加的练习(如果讲授者希望得到练习解答指导，可按照下面的网址进行查找).

在第 4、5 章描述的时间序列的小波分析能够容易地实现对于计算离散小波变换和极大重叠离散小波变换(及其逆)的基本算法. 而这些算法能够直接和容易地使用 4.6 节与 5.5 节的评论和扩展中的伪代码进行编写，可以通过本书网站链接到现有的 S-Plus 和 Lisp 中的软件. 本书网站是

http://www.staff.washington.edu/dbp/wmtsa.html

(另外，读者可以进入剑桥大学出版社网站，现在是 http://www.cup.org，然后搜索本书.)读者还可以通过这个网站得到勘误表、更新的参考文献等附加材料. 另外，读者可以通过该网站下载各种滤波器系数(如 4.8 节和 4.9 节讨论的)、书中作为例子使用的所有时间序列的值和一些能够用于检验计算机编码的计算的值. 为了方便教学和研讨班课程，网站还包括书中所有图表的 PDF 文件(请注意，这些图表的版权属于剑桥大学出版社，未经许可，不允许进一步分发或使用).

本书是使用 Donald Knuth 的优秀排版系统TeX进行排版的. 书中的插图使用 W. Hess(感谢他多年来的支持)创建的 GPL 绘图系统或 S-Plus 绘制而成，S-Plus 是 J. Chambers 与合作者开发并由 MathSoft 公司推向市场的 S 语言的商业版. 对于各种例子和图形的完成必须使用的计算技术是 S-Plus 或者 PITSSA(交互式时间序列和信号分析的基于 Lisp 的面向对象程序，是 Percival 部分地发展起来的).

感谢华盛顿大学应用物理实验室的 R. Spindel 和已故的 J. Harlett，他们提供了自由基金而启动本书的写作. 感谢美国国家科学基金(NSF)、美国国家健康研究所(NIH)、美国环保局(EPA)(通过在华盛顿大学的统计和环境国家研究中心)、美国海军研究署(ONR)和美国空军科学研究局(AFOSR)对本书写作的支持. 1998 年，我们在艾萨克·牛顿数学科学研究所(剑桥大学)从事非线性和非稳定信号处理计划期间的工作对于本书的完成至关重要；感谢在剑桥期间，英国工程和自然科学研究委员会(EPSRC)对我们之一(Percival)自始至终给予高级访问学者的支持.

感谢那些对手稿提出批评与评论或提供补充数据的人，他们是：G. Bardy、J. Bassingthwaighte、A. Bruce、M. Clyde、W. Constantine、A. Contreras Cristan、P. Craigmile、H.-Y. Gao、A. Gibbs、C. Greenhall、M. Gregg、M. Griffin、P. Guttorp、T. Horbury、M. Jensen、R. D. Martin、E. McCoy、F. McGraw、H. Mofjeld、F. Noraz、G. Raymond、P. Reinhall、S. Sardy、E. Tsakiroglou、B. Whitcher. 我们也很感谢许多研究生，他们对手稿和练习提出了非常有价值的批评并指出了许多错误. 感谢 I. Kang、I. MacLeod、K. Tanaka 和 Z. Xuelin，根据他们指出的错误我们在该印次上进行了校正. 对于一些遗留下来的错误(这在著作中是不可避免的)表示歉意，我们将乐于接受读者发现的任何错误，以便将它们列入网站上，并且在以后的印次中校正. 最后，感谢对于本项目提供大力支持的 David Tranah(剑桥大学出版社编辑)以及我们各自的家庭.

Don Percival
华盛顿大学应用物理实验室 355640 信箱
华盛顿州西雅图 98195-5640

Andrew Walden
英国帝国科技医学学院数学系
英国伦敦 SW7 2BZ



约定与记号

重要约定

(83)	83 页公式
(69a), (69b)	69 页公式
图 86	86 页图
表 109	109 页表
练习[72]	72 页嵌入的练习(答案在附录中)
练习[4.9]	第 4 章末的练习 9
$H(\cdot)$	函数
$H(f)$	$H(\cdot)$在 f 处的函数值
$\{h_l\}$	以整数 l 为下标的值的序列
h_l	序列的第 l 个值

缩写

ACS(autocorrelation sequence, 自相关序列), 15, 266, 341

ACVS(autocovariance sequence, 自协方差序列), 266

ANOVA(analysis of variance, 方差分析), 19, 67

AR(autoregressive process, 自回归过程), 268

ARFIMA (autoregressive, fractionally integrated, moving average process, 自回归、分形整合、移动平均过程), 285

CWT(continuous wavelet transform, 连续小波变换), 1, 10

dB(decibels, i. e., $10\log_{10}(\cdot)$, 分贝, 即 $10\log_{10}(\cdot)$), 73

DFBM(discrete fractional Brownian motion, 离散分形布朗运动), 279

DFT(discrete Fourier transform, 离散傅里叶变换), 22

DHM(Davies-Harte method, Davies-Harte 方法), 290

DWPT(discrete wavelet packet transform, 离散小波包变换), 206, 209

DWT(discrete wavelet transform, 离散小波变换), 1, 13, 56

ECG(electrocardiogram, 心电图), 125

EDOF(equivalent degrees of freedom, 等价自由度), 313

FBM(fractional Brownian motion, 分形布朗运动), 279

FD(fractionally differenced, 分形差分), 281

FFT(fast Fourier transform, 快速傅里叶变换), 28

FGN(fractional Gaussian noise, 分形高斯噪声), 279

GSSM(Gaussian spectral synthesis method, 高斯频谱合成方法), 291

Hz(Hertz：1 Hz=1 cycle per second，赫兹，1Hz=1 圈/秒)，48

IID(independent and identically distributed，独立同分布)，262

LA(least asymmetric，最接近对称)，107

LSE(least squares estimate or estimator，最小二乘估计)，374，378

MAD(median absolute deviation，中位数绝对偏差)，420

MODWPT (maximal overlap discrete wavelet packet transform，极大重叠离散小波包变换)，207，231

MODWT(maximal overlap discrete wavelet transform，极大重叠离散小波变换)，159

ML(maximum likelihood，极大似然)，341，361

MLE(maximum likelihood estimate or estimator，极大似然估计)，341，361

MRA(multiresolution analysis，多分辨分析)，65，461

MRC(mobile radio communications，移动无线电通信)，436

NMR(nuclear magnetic resonance，核磁共振)，420

NPES(normalized partial energy sequence，规范部分能量序列)，129，394-395

ODFT(orthonormal discrete Fourier transform，规范正交离散傅里叶变换)，41，46

OLSE(ordinary least squares estimate or estimator，常规最小二乘估计)，378

PDF(probability density function，概率密度函数)，256

PPL(pure power law，完全幂规律)，281

QMF(quadrature mirror filter，正交镜像滤波器)，75，474

RMSE(root mean square error，平方误差均值的平方根)，364，436

RV(random variable，随机变量)，256

SDF(spectral density function，谱密度函数)，267

SURE(Stein's unbiased risk estimator，Stein 无偏风险估计)，404

WLSE(weighted least squares estimate or estimator，加权最小二乘估计)，374

WP(wavelet packet，小波包)，209

常用非希腊字母记号

A_j [对$\{W_{j,t}\}$的平方谱密度函数 $S_j^2(\cdot)$的积分]，307

$\mathcal{A}_j$ [$N_j \times N_{j-1}$ 矩阵(行是$\{g_l\}$周期化为 N_{j-1}得到)]，94

$\widetilde{\mathcal{A}}_j$ [$N \times N$ 矩阵(行是上采样$\{\tilde{g}_l\}$周期化为 N 得到)]，176

$\mathcal{A}_L(\cdot)$ [关于$\mathcal{H}^{(D)}(\cdot)$的低通成分的平方增益函数]，106

$\{a_{n,t}\}$ [第 n 个正弦锥]，274

$\arg(z)$ [复数 z 的辐值]，21

B [后移算子]，283

B_t [离散分形布朗运动]，279

$B_H(\cdot)$ [分形布朗运动]，279

$\mathcal{B}_j$ [$N_j \times N_{j-1}$ 矩阵(行是$\{h_l\}$周期化为 N_{j-1}得到)]，94

$\widetilde{\mathcal{B}}_j$ [$N \times N$ 矩阵(行由上采样$\tilde{h}_l$周期化为 N 得到)]，176

C_j [在倍频程带$\left[\frac{1}{2^{j+1}}, \frac{1}{2^j}\right]$上的谱密度函数 $S_X(\cdot)$的平均值]，343

$\widetilde{C}_j$ [C_j 对分形差分过程的近似]，344

$\{C_n\}$ [规范部分能量序列]，129，395

$\boldsymbol{C}$ [N 维统计随机随机向量]，393

$\mathrm{cov}\{\cdot,\cdot\}$ [协方差算子]，259

$\mathcal{D}(\cdot)$ [微分滤波器的平方增益函数]，105

$\mathcal{D}_j$ [第 j 层小波细节(DWT)]，64

$\widetilde{\mathcal{D}}_j$ [第 j 层小波细节(MODWT)]，169，171

$\mathbf{D}$ [在匹配追踪中使用的字典(向量集)]，239

$\boldsymbol{D}$ [N 维确定信号向量]，393

d [差分运算数]，287

$d_{j,t}$ [离散小波变换 $\boldsymbol{d}$ 第 j 层的第 t 分量]，419

d_l [向量 $\boldsymbol{d}$ 的第 l 个分量]，398

$\boldsymbol{d}$ [对确定信号 $\boldsymbol{D}$ 的变换系数]，398

$\boldsymbol{d}_\gamma$ [字典元素(在匹配追踪字典 $\mathbf{D}$ 中的向量)]，239

$E\{\cdot\}$ [期望算子]，256，258

$E\{X_0 \mid X_1=x_1\}$ [已知 $X_1=x_1$，X_0 的条件期望]，260

ε_X [向量 $\boldsymbol{X}$ 的能量(2-范数)]，42，72

e [2.718 281 828 459 045]，3，21

e^{ix} [复指数]，21

$e_{j,t}$ [$\boldsymbol{e}$ 的离散小波变换第 j 层的第 t 个分量]，419

e_l [向量 $\boldsymbol{e}$ 的第 l 个分量]，398

$\boldsymbol{e}$ [独立同分布噪声 $\boldsymbol{\varepsilon}$ 的变换系数]，398

$\{F_k\}$ [规范正交离散傅里叶变换]，46

$\mathcal{F}$ [$N\times N$ 规范正交离散傅里叶变换矩阵]，47

$\boldsymbol{F}$ [规范正交傅里叶变换系数$\{F_k\}$形成的向量]，47

f [正弦频率]，22

f_k [k/N 或 $k/(N\Delta t)$，第 k 个傅里叶频率]，28，87

f_N [奈奎斯特频率]，87，267

$f_X(\cdot)$ [随机变量 X 的概率密度函数]，256

$f_{X_0,X_1}(\cdot,\cdot)$ [随机变量 X_0 和 X_1 的联合概率密度函数]，258

$f_{X_0 \mid X_1=x_1}(\cdot)$ [已知 $X_1=x_1$，随机变量 X_0 的条件概率密度函数]，260

$G(\cdot)$ [$\{g_l\}$的传递函数]，76，154

$\widetilde{G}(\cdot)$ [$\{\tilde{g}_l\}$的传递函数]，163，202

$G_j(\cdot)$ [在 $G_1(\cdot)\equiv G(\cdot)$时，$\{g_{j,l}\}$的传递函数]，97，154

$\widetilde{G}_j(\cdot)$ [在$\widetilde{G}_1(\cdot)\equiv\widetilde{G}(\cdot)$时，$\{\tilde{g}_{j,l}\}$的传递函数]，169，202

$\mathcal{G}(\cdot)$ [$\{g_l\}$的平方增益函数]，76，154

$\widetilde{\mathcal{G}}(\cdot)$ [$\{\tilde{g}_l\}$的平方增益函数]，163，202

$\mathcal{G}_j(\cdot)$ [在$\mathcal{G}_1(\cdot)\equiv\mathcal{G}(\cdot)$，$\{g_{j,l}\}$的平方增益函数]，154

$\widetilde{\mathcal{G}}_j(\cdot)$ [在$\widetilde{\mathcal{G}}_1(\cdot)\equiv\widetilde{\mathcal{G}}(\cdot)$，$\{\tilde{g}_{j,l}\}$的平方增益函数]，202

$\mathcal{G}^{(D)}(\cdot)$ [Daubechies 尺度滤波器$\{g_l\}$的平方增益函数]，105

$\{g_l\}$ [离散小波变换尺度滤波器]，75，154，463

$\{\tilde{g}_l\}$ [极大重叠离散小波变换尺度滤波器]，163，202

$\{g_l^\circ\}$ [$\{g_l\}$周期化为 N 得到]，77

$\{\tilde{g}_l^\circ\}$ [$\{\tilde{g}_l\}$周期化为 N 得到]，168

$\{\overline{g}_l\}$ [翻转尺度滤波器，例如：$\overline{g}_l=g_{-l}$]，463

$\{g_l^{(ep)}\}$ [极小相位 Daubechies 尺度滤波器]，106

$\{g_l^{(la)}\}$ [最接近对称 Daubechies 尺度滤波器]，107

$\{g_{j,l}\}$ [第 j 层离散小波变换尺度滤波器，$\{g_{1,l}\}\equiv\{g_l\}$]，96，154

$\{\tilde{g}_{j,l}\}$ [第 j 层极大重叠离散小波变换尺度滤波器，$\{\tilde{g}_{1,l}\}\equiv\{\tilde{g}_l\}$]，169，202

$\{g_{j,l}^\circ\}$ [$\{g_{j,l}\}$周期化为 N 得到]，97

$\{\tilde{g}_{j,l}^\circ\}$ [$\{\tilde{g}_{j,l}\}$周期化为 N 得到]，170

$P[A]$ [事件 A 发生的概率]，256

$Q_\eta(p)$ [χ^2_η 分布的 $p\times 100\%$ 的百分点]，263-264

$R_{j,t}$ [$\boldsymbol{R}$ 的离散小波变换的 j 层第 t 个分量]，424

R_l [向量 $\boldsymbol{R}$ 的第 l 个分量]，407

$\mathcal{R}_j$ [第 j 层小波粗糙(DWT)]，66

$\mathbb{R}$ [整个实轴]，457

$\mathbb{R}^N$ [实值 N 维向量空间]，45

$\mathcal{R}(z)$ [复数 z 的实部]，21

$\boldsymbol{R}$ [随机信号 $\boldsymbol{C}$ 的变换系数]，407

$S_j(\cdot)$ [$\{\widetilde{W}_{j,t}\}$ 或 $\{W_{j,t}\}$ 非边界部分的谱密度函数]，304，348

$S_X(\cdot)$ [(能量)谱密度函数]，267

$\hat{S}_X^{(mt)}(\cdot)$ [多锥谱密度函数估计]，274

$\hat{S}_{X,n}^{(mt)}(\cdot)$ [用于形成 $\hat{S}_X^{(mt)}(\cdot)$ 的第 n 个特征谱]，274

$\hat{S}_X^{(p)}(\cdot)$ [周期图]，269

$\mathcal{S}_j$ [离散小波变换的第 j 层小波光滑]，64

$\widetilde{\mathcal{S}}_{J_0}$ [极大重叠离散小波变换的第 J_0 层光滑]，169，171

$\{s_{X,\tau}\}$ [自协方差序列]，266

$\{\hat{s}_{X,\tau}^{(p)}\}$ [自协方差序列的有偏估计]，269

$\mathcal{T}$ [$N\times N$ 循环位移矩阵或单位延迟算子]，52，457

t [(连续)实际时间或(离散)非单位指标]，5，24

$U_{j,n}(\cdot)$ [$\{u_{j,n,l}\}$的传递函数]，215

$\widetilde{U}_{j,n}(\cdot)$ [$\{\tilde{u}_{j,n,l}\}$的传递函数]，232

$\{u_{j,n,l}\}$ [节点$(j，n)$的离散小波包变换的滤波器]，214

$\{\tilde{u}_{j,n,l}\}$ [节点$(j，n)$的极大重叠离散小波包变换的滤波器]，231

V_j [尺度 λ_j 的函数的近似子空间]，462

$V_{j,t}$ [$\boldsymbol{V}_j$ 的第 t 个元素]，94

$\widetilde{V}_{j,t}$ [$\widetilde{\boldsymbol{V}}_j$ 的第 t 个元素]，169

$\mathcal{V}_j$ [$\boldsymbol{X}$ 到 $\boldsymbol{V}_j$ 的 $N_j\times N$ 映射矩阵]，94

$\widetilde{\mathcal{V}}_j$ [$\boldsymbol{X}$ 到 $\widetilde{\boldsymbol{V}}_j$ 的 $N\times N$ 映射矩阵]，171

$\boldsymbol{V}_j$ [第 j 层离散小波变换尺度系数向量]，94

$\widetilde{\boldsymbol{V}}_j$ [第 j 层极大重叠离散小波变换尺度系数向量]，169

var{·} [方差算子]，259

W_j [尺度为 τ_j 的函数细节子空间]，472

$W_{j,n,t}$ [$\boldsymbol{W}_{j,n}$的第 t 个元素]，214

$\widetilde{W}_{j,n,t}$ [$\widetilde{\boldsymbol{W}}_{j,n}$的第 t 个元素]，231

$W_{j,t}$ [$\boldsymbol{W}_j$ 的第 t 个元素]，94

$\widetilde{W}_{j,t}$ [$\widetilde{\boldsymbol{W}}_j$ 的第 t 个元素]，169

$\{\widetilde{W}_{j,t}\}$ [随机过程$\{X_t\}$的第 j 层极大重叠离散小波变换系数]，296

W_n [第 n 个离散小波变换系数]，57

$\mathcal{W}$ [$N\times N$ 离散小波变换矩阵]，57

$\mathcal{W}_j$ [$\boldsymbol{X}$ 到 $\boldsymbol{W}_j$ 的 $N_j\times N$ 映射矩阵]，94

$\widetilde{\mathcal{W}}_j$ [$\boldsymbol{X}$ 到 $\widetilde{\boldsymbol{W}}_j$ 的 $N\times N$ 映射矩阵]，171

$\boldsymbol{W}$ [包含离散小波变换系数$\{W_n\}$的向量]，57，150

$\boldsymbol{W}_j$ [第 j 层离散小波变换小波系数向量]，94

$\widetilde{\boldsymbol{W}}_j$ [第 j 层极大重叠离散小波变换小波系数向量]，169

$\boldsymbol{W}_{j,n}$ [结点$(j，n)$处的离散小波包变换系数向量]，209

$\widetilde{\boldsymbol{W}}_{j,n}$ [结点$(j，n)$处的极大重叠离散小波包变换系数向量]，232

width$_a${·} [自相关宽度]，12，103

$X_0，\cdots，X_{N-1}$ [随机过程的时间序列或划分]，41，269

$\{X_t\}$ [时间序列或随机过程]，41，266，295-296

$\overline{X}$ [$X_0，\cdots，X_{N-1}$ 的采样均值(算术平均)]，48

$\overline{X}_t(\lambda)$ [$X_{t-\lambda+1}，X_{t-\lambda+2}，\cdots，X_t$ 的采样均值]，58

$\{\mathcal{X}_k\}$ [$\{X_t\}$的离散傅里叶变换], 72

$\mathbf{X}$ [X_0, …, X_{N-1}形成的向量], 41-42

$Y^{(mt)}(f)$ [对数多锥谱密度函数估计外加常数], 276

$Y^{(p)}(f)$ [对数周期图外加常数], 271

Z [单位均值零方差的高斯正态随机变量], 257

常用希腊字母记号

α [幂率谱密度函数的指数], 279, 281, 286

α [检验的重要水平], 373, 434

β [与分形差分参数δ有关的线性回归模型中的斜率], 374

$\hat{\beta}^{(wls)}$ [β的加权最小二乘估计], 376

$\Gamma(\cdot)$ [伽马函数], 257

γ [匹配追踪字典中的向量指标], 239

γ [欧拉常数(0.577 215 664 901 532…)], 270, 432

$\gamma_{J_0}^{(G)}$, $\gamma_j^{(H)}$ [$\mathbf{V}_{J_0}$, $\mathbf{W}_j$中的系数指标], 137,147

$\gamma_{J_0}^{(G)}$, $\gamma_j^{(H)}$ [$\mathbf{V}_{J_0}$, $\mathbf{W}_j$"早期"边界系数的个数], 137, 147

γ_l^2 [高斯混合模型中分量方差比], 410

$\gamma(\cdot)$ [实值函数], 457

$\gamma_{j,k}(\cdot)$ [$\gamma(\cdot)$的平移和伸缩], 459

Δt [采样间隔], 48, 59

δ [普通阈值], 223, 399

δ [分形差分过程的长记忆参数], 283-284, 286, 288

$\delta^{(s)}$ [稳定分形差分过程的长记忆参量], 288, 368

$\delta^{(S)}$ [基于 Stein 无偏有损估计的阈值], 405

$\delta^{(u)}$ [统一阈值], 400

$\delta_{j,k}$ [Kronecker δ函数], 42-43

$\hat{\delta}$ [稳定分形差分过程的δ的恰当极大似然估计], 368

$\hat{\delta}^{(loocv)}$ [漏一交叉阈值], 402, 423

$\hat{\delta}^{(tfcv)}$ [二重交叉阈值], 402, 422

$\hat{\delta}^{(wls)}$ [稳定或非稳定分形差分过程的δ加权最小二乘估计], 377

$\tilde{\delta}^{(s)}$ [稳定分形差分过程的δ近似加权最小二乘估计], 363

$\tilde{\delta}^{(s/ns)}$ [如：$\tilde{\delta}^{(s)}$，但是针对非稳定分形差分过程], 371

$\boldsymbol{\varepsilon}$ [独立同分布随机变量形成的N维向量], 393

ε [小的正数], 2, 486

$\varepsilon(f)$, $\varepsilon(f_k)$ [频域模型(非相关)的误差项], 270, 432

ε_t [非相关随机变量(白噪声)序列的第t项], 268

ζ [线性回归模型的截距], 374

$\boldsymbol{\eta}$ [非独立同分布随机变量形成的N维向量], 393

η [χ^2分布的自由度], 263, 313

η_1, η_2, η_3, η_j [小波方差估计的等价自由度], 313-4, 376

$\eta(f)$, $\eta(f_k)$ [频域模型(相关)的误差项], 276, 440

θ [极坐标表示$z=|z|e^{i\theta}$中的幅角], 21

θ [贝叶斯模型中先验分布的参数], 264

$\theta(\cdot)$ [滤波器的相位函数], 25

$\theta^{(G)}(\cdot)$ [离散小波变换尺度滤波器的相位函数], 106

$\theta^{(H)}(\cdot)$ [离散小波变换小波滤波器的相位函数], 112

$\theta_{c_j,n,m}(\cdot)$ [离散小波包变换小波滤波器的相位函数的分量], 229

ϑ [t分布的自由度], 257, 426

κ [规范t分布的尺度参数], 258, 414

κ [一致分布在整数0, 1, …, $N-1$的随机变量], 356

Λ_X [$N \times N$ 对角协方差矩阵]，355

λ [尺度(区间或平均长度)]，6，58

λ_j [2^j，$j(j \geq 1)$层尺度函数的非单位尺度]，85，481

μ [随机变量的期望值]，256-257

ν [频域自协方差的延迟]，276-277

ν [时间序列或滤波器的前移]，111-112

$\nu_j^{(G)}$、$\nu_j^{(H)}$ [尺度滤波器，小波滤波器的前移]，114

$\nu_{j,n}$ [小波包滤波器的前移]，229

$\nu_X^2(\tau_j)$ [在尺度 τ_j 的小波方差]，296

$\hat{\nu}_X^2(\tau_j)$ [在尺度 τ_j 的小波方差的无偏极大重叠离散小波变换估计]，306

$\hat{\hat{\nu}}_X^2(\tau_j)$ [在尺度 τ_j 的小波方差的无偏离散小波变换估计]，308

$\tilde{\nu}_X^2(\tau_j)$ [在尺度 τ_j 的小波方差的有偏极大重叠离散小波变换估计]，306

$\tilde{\tilde{\nu}}_X^2(\tau_j)$ [在尺度 τ_j 的小波方差的有偏离散小波变换估计]，308

π [3.141 592 653 589 793…]，3，21

ρ [两个随机变量的相关性]，259

$\rho_{X,\tau}$ [稳定过程在延迟 τ 的自相关序列]，266

$\hat{\rho}_{X,\tau}$ [在延迟 τ 的自相关序列的估计]，16，341

Σ_X [随机向量 $\boldsymbol{X}$ 的自方差矩阵]，259，262

$\tilde{\Sigma}_X$ [自方差矩阵 Σ_X 的基于小波的近似]，362

σ^2，σ_X^2 [随机变量的方差]，3，257，279

σ_ϵ^2 [独立同分布过程的方差]，393

σ_ϵ^2 [白噪声过程的方差]，268

$\sigma_{G_l}^2$ [高斯混合模型的部分的方差]，410

$\sigma_{n_l}^2$ [非独立同分布过程的方差]，403

$\hat{\sigma}_X^2$，$\hat{\sigma}_Y^2$ [采样均值得到的采样方差]，48，299

$\tilde{\sigma}_Y^2$ [过程均值得到的采样方差]，299

$\hat{\sigma}_{(mad)}^2$ [中位数绝对偏差得到的方差估计]，420

$\tilde{\sigma}_{(mad)}^2$ [如$\hat{\sigma}_{(mad)}^2$，但是基于极大重叠离散小波变换的]，429

τ [自相关或自协方差序列的延迟指标]，16，266

τ_j [2^{j-1}，j 层小波系数的非单位尺度($j \geq 1$)]，59

$\Upsilon(\cdot)$ [最小化该函数得到 Stein 无偏风险估计阈值 $\delta^{(S)}$]，405

υ，υ_l [拉普拉斯分布的逆方差]，257，265，413

$\phi(\cdot)$ [尺度函数]，459

$\phi_{j,k}(\cdot)$ [尺度函数的伸缩平移]，460

$\phi^{(H)}(\cdot)$ [哈尔尺度函数]，460

$\phi_{p,1}$，…，$\phi_{p,p}$ [自回归过程(p)的系数]，268，292

χ_η^2 [自由度为 η 的 χ^2 随机变量]，263

$\psi(\cdot)$ [小波函数]，2，474

$\psi(\cdot)$，$\psi'(\cdot)$ [Γ_2，Γ_3 函数]，275，376，440

$\psi_{j,k}(\cdot)$ [小波函数的伸缩平移]，474

$\psi^{(H)}(\cdot)$ [哈尔小波函数]，2，475

$\psi^{(Mh)}(\cdot)$ [墨西哥帽小波函数]，3

ω_0 [Morlet 小波函数的参数]，4

其他数学记号和符号

$\approx$ [约等于]，83-84，264，297

$\{a \star a_t\}$ [实值序列$\{a_t\}$的自相关]，69

$\{a^* \star a_t\}$ [复值序列$\{a_t\}$的自相关]，25，30，36-7

$\mathcal{O}^H$ [矩阵的复共轭转置]，45

z^* [z 的复共轭]，21

$\{a^* \star b_t\}$ [序列$\{a_t\}$和$\{a_t\}$的复互相关]，24-25，30

$\in$ [属于]，2，398

$\overline{\in}$ [不属于]，2，398

$\{a * b_t\}$ [序列$\{a_t\}$，$\{b_t\}$的卷积]，24，30，36-37

$|\Sigma_X|$ [矩阵 Σ_X 的行列式]，361

$\downarrow$ [下抽样]，70，80，92，96

$\doteq$ [一定精度下相等(如，$\pi \doteq 3.14$ 或 $\pi \doteq 3.141\,6$)]，3，73

$\equiv$ [定义相等]，20

$\stackrel{d}{=}$ [分布意义下的相等]，257

$\hat{\cdot}$ [估计，例如：$\hat{\nu}_X^2(\tau_j)$ 是 $\nu_X^2(\tau_j)$ 的估计]，306

$\{a_t\} \leftrightarrow \{A_k\}$ [傅里叶变换对($\{a_t\}$是有限序列)]，29，36

$\{a_t\} \leftrightarrow A(\cdot)$ [傅里叶变换对($\{a_t\}$是无穷序列)]，23，35

$\lfloor x \rfloor$ [$\leqslant x$ 的最大整数]，50，146

$\lceil x \rceil$ [$\geqslant x$ 的最小整数]，50，146

$1_{\mathcal{J}}(\cdot)$ [集$\mathcal{J}$的特征函数]，404

$\langle \cdot, \cdot \rangle$ [内积]，42，45

$|z|$ [z 的模]，21

$\mathbf{1}$ [1 的 N 维向量]，50

$\{a^{\circ}_t\}$ [无穷序列$\{a_t\}$的周期化(长为 N)]，33

$[a, b]$ [满足 $a \leqslant x \leqslant b$ 的所有 x 的集合]，22

(a, b) [满足$a<x<b$的所有x的集合]，2

$(a, b]$ [满足 $a < x \leqslant b$ 的所有 x 的集合]，465

$\|\cdot\|^2$ [平方范数]，42，46

$\mathbf{X}^T$ [向量 $\mathbf{X}$ 的转置]，42

$\mathcal{O}^T$ [矩阵$\mathcal{O}$的转置]，42

$\uparrow$ [上抽样]，82，95，201

$\mathcal{O}_{j\cdot}$ [$N \times N$ 矩阵$\mathcal{O}$的第 j 行向量]，42

$\mathcal{O}_{\cdot k}$ [$N \times N$ 矩阵$\mathcal{O}$的第 k 列向量]，42

0，$\mathbf{0}$， [零向量]，101

目　录

第1章

小波导引

1.0 引言

小波是分析时间序列或图像的数学工具(不排除在其他领域的应用，例如，参见 Stollnitz 等(1996)和 Sweldens(1996))．本书中，我们仅仅关注小波在时间序列分析方面的应用，这里的时间序列是指和一个有序的独立变量 t 有关的任何观察序列(变量 t 可以取离散值，例如整数；或者取连续值，例如整个实轴上的值——两类例子包括沿一条直线的时间、深度或者距离，也就是说时间序列并非必须与时间有关)．小波是分析时间序列的一种相当新颖的方法，其正式的研究可追溯到 20 世纪 80 年代，但是在许多方面，小波仅仅是具有新的优美的数学结果和有效的计算方法的古老概念的合成．在某些情况下，小波分析是现存的分析技术(例如，相关分析与谱分析)的补充，而在其他情况下可以用来解决一些在小波引入之前进展缓慢的问题．

粗略地说(用不尽贴切的相关语!)，在小波的发展过程中主要有两种"波"．第一种"波"导致了所谓的连续小波变换(CWT)，它可以设计成用来处理定义在整个实轴上的时间序列；第二种"波"即离散小波变换(DWT)中的波，可以用来处理定义在一定范围内的整数点上的时间序列(通常 $t=0, 1, 2, \cdots, N-1$，N 表示时间序列的个数)．在这一章中我们通过连续小波变换来介绍小波．重点是表达小波分析的思想而不是详细介绍它的数学发展，这在许多其他地方也是可取的．我们的方法将集中于这样一种情况，即小波分析到底可以告诉我们关于时间序列的什么内容．我们不推测它和其他常用的分析技术(如傅里叶分析)的广泛的相似性．在 1.1 节和 1.2 节后，将比较离散小波变换和连续小波变换，并且讨论为什么我们觉得离散小波变换是离散时间序列分析的一个自然工具．本书余下的部分将从基础开始介绍离散小波变换和某些密切相关的变换(从第 2 章至第 6 章)，然后在第 7 章至第 10 章通过离散小波变换讨论时间序列的统计分析．在

第11章中我们回到连续小波变换，在这一章中我们通过定义在整个实轴上的函数的优美的多分辨分析理论说明离散小波变换和连续小波变换的联系，以此加深对离散小波变换的理解.

1.1 小波的本质

什么是小波？正如名字所暗示的，小波是“小的波”．一个小的波本质上生成和衰减于一个有限的时间周期内．相对的概念显然是“大的波”．大波的一个例子是正弦函数，$\sin(u)$ 的图形在 $u\in(-\infty,\infty)$ 上保持上下振荡．为了开始量化小波的概念，让我们先来考虑定义于整个实轴 $(-\infty,\infty)$ 上的满足下面两个基本性质的实值函数 $\psi(\cdot)$：

[1]$\psi(\cdot)$的积分为零：

$$\int_{-\infty}^{\infty}\psi(u)\mathrm{d}u=0.\tag{2a}$$[○]

[2]$\psi(\cdot)$的平方的积分为1：

$$\int_{-\infty}^{\infty}\psi^2(u)\mathrm{d}u=1.\tag{2b}$$

(对于正弦函数，上面的积分是无限的，故 $\sin^2(\cdot)$不能重新规范化使其积分为1.)

如果等式(2b)成立，那么对于任何满足 $0<\epsilon<1$ 的 ϵ，一定存在一个区间$[-T,T]$，使

$$\int_{-T}^{T}\psi^2(u)\mathrm{d}u>1-\epsilon.$$

如果我们把 ϵ 看作非常接近于0，那么 $\psi(\cdot)$在区间 $[-T,T]$ 外仅仅是不显著地偏离零．它的非零变动范围主要限制于一个有限的区间 $[-T,T]$．因为区间 $[-T,T]$ 的长度和整个实轴的长度比较起来非常小，因此可以认为 $\psi(\cdot)$的非零变动范围限制在一个相对较小的时间区间上．等式(2b)说明 $\psi(\cdot)$相对于零必定有一些偏离，然而等式(2a)告诉我们大于零的偏移必须被小于零的偏移所抵消，因此 $\psi(\cdot)$必然像一个波．所以等式(2a)和(2b)导致了“小的波”或者小波.

图3绘出了三个这样的小波，根据下面的定义，读者可以验证这些函数确实满足式(2a)和(2b)(积分为零可以从图中看出)．第一个小波称为哈尔(Haar)小波函数：

$$\psi^{(\mathrm{H})}(u)\equiv\begin{cases}-1/\sqrt{2}, & -1<u\leqslant 0;\\ 1/\sqrt{2}, & 0<u\leqslant 1;\\ 0, & \text{其他}.\end{cases}\tag{2c}$$

(和这个小波有一些细微差别的公式将在11.6节中详细讨论)．这个小波可以说

○ 为方便读者查阅原书网站的相关资料，本书序号严格按照原书来编排．——编辑注

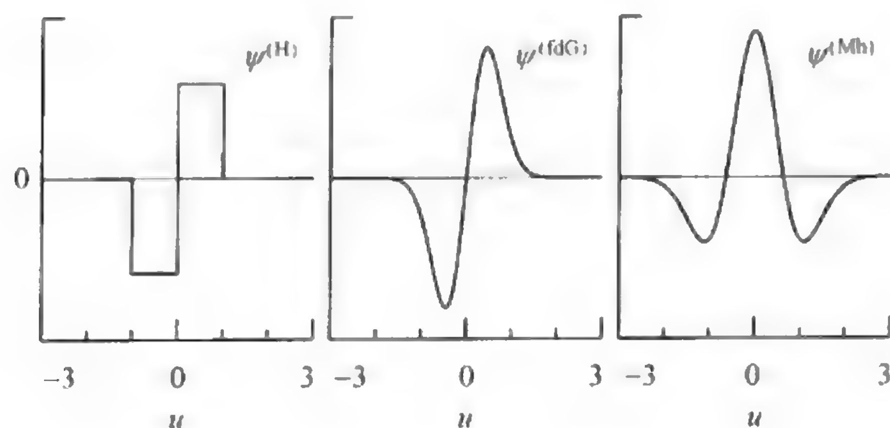


图 3 三种小波. 从左到右分别是一种哈尔小波、有关高斯概率密度函数(PDF)的一阶导数的小波、墨西哥帽小波(有关高斯概率密度函数的二阶导数的小波)

是最古老的小波，是以哈尔(A. Haar)的名字命名的，他在 1910 年的一篇论文中将它发展为一种分析工具. 为了形成另外两种小波，我们从高斯(正态)概率密度函数

$$\phi(u) \equiv \frac{e^{-u^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}, \quad -\infty < u < \infty$$

出发，其均值为 0，方差为 σ^2. $\phi(\cdot)$的一阶导数是

$$\frac{d\phi(u)}{du} = -\frac{ue^{-u^2/2\sigma^2}}{\sigma^3\sqrt{2\pi}}.$$

如果将上面的函数的负函数再作规范化，使其满足式(2b)，就得到小波

$$\psi^{(fdG)}(u) \equiv \frac{\sqrt{2}ue^{-u^2/2\sigma^2}}{\sigma^{3/2}\pi^{1/4}}, \tag{3a}$$

如图 3 中间所示，$\sigma \doteq 0.443\,11$. 再次利用适当的规范化，$\phi(\cdot)$的二阶导数的负值也产生一个小波，通常称为墨西哥帽小波，如下所示：

$$\psi^{(Mh)}(u) \equiv \frac{2\left(1-\frac{u^2}{\sigma^2}\right)e^{-u^2/2\sigma^2}}{\pi^{1/4}\sqrt{3\sigma}}. \tag{3b}$$

这个小波的名字来源于其小波函数图像(图 3 中右边的图像，$\sigma \doteq 0.636\,28$)的形状.

总之，小波被定义为积分为零且平方可积的函数(另可参见下面的评论与扩展的条目[1]).

1.1节的评论与扩展

[1]我们已经给出了小波的一个"骨架"定义，这样就可以阐述这一主题背后的关键概念. 为了得到实用的小波，我们必须在式(2a)和(2b)的基础上再加一些条件. 关于小波大量的数学研究表明，为了达到用小波分析一些对象(例如边缘检测、奇异性分析等)的目的，必须确定需要什么条件. 这里只是提一下一个重要而普通的条件，即所谓的容许性条件. 一个小波 $\psi(\cdot)$ 称为是容许的，如果它的傅里叶变换，即

$$\Psi(f) \equiv \int_{-\infty}^{\infty} \psi(u) \mathrm{e}^{-\mathrm{i}2\pi f u} \mathrm{d}u,$$

使得

$$C_\psi \equiv \int_0^{\infty} \frac{|\Psi(f)|^2}{f} \mathrm{d}f \quad 满足\ 0 < C_\psi < \infty \tag{4a}$$

(第 2 章将给出傅里叶理论的一个综述，包括在 2.7 节后，对像 $\Psi(\cdot)$ 这样的函数的一些关键结果的小结). 这个条件保证了从函数 $x(\cdot)$ 的连续小波变换可以重构这个函数（参见式(11a)). 关于容许性条件的其他论述可参考 Daubechies(1992, pp. 24-26).

[2]为了简化本章主要部分的阐述，我们假定 $\psi(\cdot)$ 是实值的，但是也经常要用到复值小波，特别是在地球物理应用中(例如，参看 Foufoula-Georgiou and Kumar，1994). 小波分析的第一篇文献(Goupillaud 等，1984)是由 Morlet 在做石油和天然气勘探过程中分析地球物理信号引发的，他想分析既包含少量周期的短时高频瞬变又包含长低频瞬变的信号. 文中的例子用复小波

$$\psi(u) = C\mathrm{e}^{-\mathrm{i}\omega_0 u}\left(\mathrm{e}^{-u^2/2} - \sqrt{2}\mathrm{e}^{-\omega_0^2/4}\mathrm{e}^{-u^2}\right) \tag{4b}$$

给出，其中，C 和 ω_0 是常数. 众所周知，

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}\omega_0 u}\mathrm{e}^{-u^2/2}\mathrm{d}u = \sqrt{2\pi}\mathrm{e}^{-\omega_0^2/2}$$

(参见 Bracewell，1978，p. 386 或者 Percival and Walden，1993，p. 67)，由此得到 $\int_{-\infty}^{\infty} \psi(u)\mathrm{d}u = 0$，因此式(2a)满足. 对 Morlet 来说，式(2a)成立的事实也许并不只是一个数学结果，还有物理上的必然：分析中的地震反射时间序列也"积分为零"，因为压缩和变稀薄必须相互抵消. 常数 C 的选择使得等式(2b)的复值形式成立，即 $\int_{-\infty}^{\infty} |\psi(u)|^2 \mathrm{d}u = 1$ 对于一个特别选择的 ω_0 成立. 例如，当 $\omega_0 = 5$ 时，我们有 $C \doteq 0.7528$. 随着 ω_0 的进一步增加，式(4b)的负项变得可以忽略；当 $\omega_0 = 10$ 而 $C = \pi^{-1/4}$ 时，我们有 $\int_{-\infty}^{\infty} |\psi(u)|^2 \mathrm{d}u \doteq 1$ 可以精确到小数点后第九位. 因此，对于一个相对大的 ω_0，

$$\psi(u) \approx \psi_{\omega_0}^{(\mathrm{M})}(u) \equiv \pi^{-1/4}\mathrm{e}^{-\mathrm{i}\omega_0 u}\mathrm{e}^{-u^2/2},$$

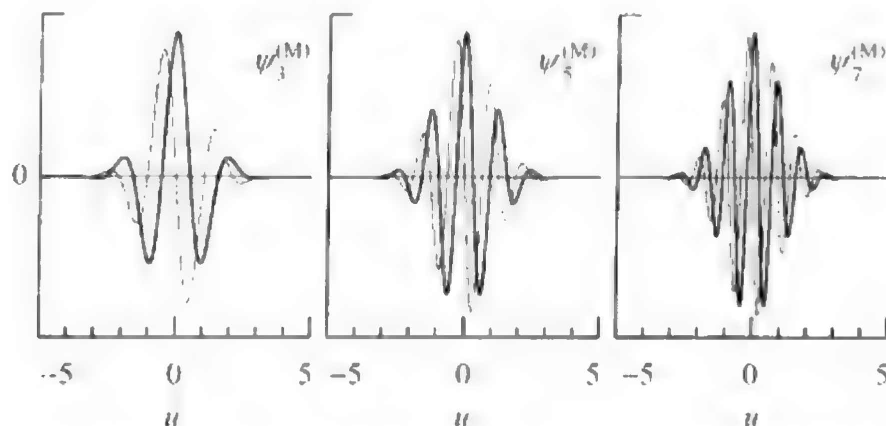


图 5　三个 Morlet 小波 $\psi_{\omega_0}^{(M)}(\cdot)$. 这些小波是复值小波，所以它们的实部和虚部分别用粗线和细线画在图像中. 参数ω_0控制复指数函数的频率，而复指数函数由一个形状由标准高斯概率密度函数所确定的函数调节. 当ω_0从 3 增加到 7 时，高斯概率密度函数在有效宽度内的振荡次数增加

这通常称为 Morlet 小波(参见 Kumar and Foufoula-Georgiou, 1994). 然而，我们注意到由于 $\psi(\cdot)$的确积分为零，$\psi_{\omega_0}^{(M)}(\cdot)$不能也为 0. 对于$\omega_0=3$，5 和 7，$\psi_{\omega_0}^{(M)}(\cdot)$的图像在图 5 中给出.

Morlet 小波本质上是频率 $f_0=\omega_0/(2\pi)$的复指数，它的幅度由一个与标准高斯概率密度函数成比例的函数调节. 正如我们在下一节中讨论的，像图 3 中的小波提供函数在不同尺度上平均变化的局部分析，而 Morlet 小波生成的分析却不能直接用不同尺度下的平均变化来解释(然而，Morlet 小波和图 3 中的小波在频域上是十分相似的——都是一个带通滤波器).

1.2　小波分析的本质

我们已经有了小波的定义，并且有了一些直观的概念，但是它有什么用处呢？一个立即的答案是，像图 3 中所示的小波可以告诉我们，某些其他函数的加权平均是如何从一个平均周期变化到另一个周期的. 小波分析的这个解释是一个重要的概念，所以我们在本节中详细讨论. 开始，我们设 $x(\cdot)$是关于独立变量 t 的实值函数，这个变量称为“时间”(这仅仅是为了方便，实际上，例如，t 可以具有深度单位而不是时间单位). 我们将 $x(\cdot)$称为一个“信号”也仅仅是为了方便. 考虑积分

$$\frac{1}{b-a}\int_a^b x(u)\,\mathrm{d}u \equiv \alpha(a,b), \tag{5}$$

这里假设 $a<b$，$x(\cdot)$使得上面的积分有意义. 在初等微积分中，$\alpha(a,b)$是 $x(\cdot)$在区间$[a, b]$上的平均值，它和 N 个观察值集合的采样平均的概念有关. 为了理解这个关系，假设 $x(\cdot)$是如下形式的阶梯函数：

$$x(t)=x_j,\quad a+\frac{j}{N}(b-a)<t\leqslant a+\frac{j+1}{N}(b-a),\quad j=0,1,2,\cdots,N-1,$$

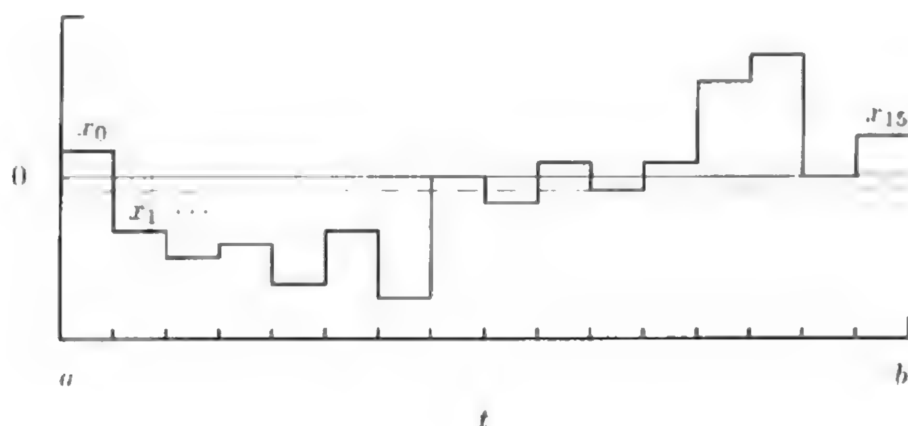


图6 阶梯函数 $x(\cdot)$在16等分区间 $[a, b]$ 形成的子区间上相继取值 x_0, x_1, …, x_{15}. 根据式(5)的定义, $x(\cdot)$在 $[a, b]$上的均值恰恰是所有16个 x_j的抽样平均(图中的虚线)

(参见图6.)由黎曼(Riemann)积分的定义容易得到

$$\frac{1}{b-a}\int_a^b x(u)\mathrm{d}u = \frac{1}{b-a}\sum_{j=0}^{N-1} x_j \frac{b-a}{N} = \frac{1}{N}\sum_{j=0}^{N-1} x_j.$$

除了将均值 $a(a, b)$ 作为区间$[a, b]$端点的函数之外，我们很容易将其视为区间长度 $\lambda \equiv b-a$ 和区间的时间中心 $t=(a+b)/2$ 的函数. λ 称为关于平均的尺度. 通过 λ 和 t 能够定义

$$A(\lambda, t) \equiv a\left(t-\frac{\lambda}{2}, t+\frac{\lambda}{2}\right) = \frac{1}{\lambda}\int_{t-\frac{\lambda}{2}}^{t+\frac{\lambda}{2}} x(u)\mathrm{d}u.$$

我们称 $A(\lambda, t)$ 为信号 $x(\cdot)$在尺度 λ 上关于时间中心 t 的均值.

在自然科学中，信号在不同尺度上的均值是非常有意义的，在一些科学文献分析中最近研究的例子包括：

[1]加拿大 Ontario 湖落叶林中和森林上空的温度和垂直速度的二分之一平均(Gao and Li, 1993).

[2]俄克拉何马州 Norman 上空每十分钟和每小时强风暴下的降水量(Kumar and Foufoula-Georgiou, 1993).

[3]沿太平洋赤道的月平均海平面温度(Mak 等, 1995).

[4]大尺度范围的海上层积云的非匀质性(Gollmer 等, 1995).

[5]30分钟平均垂直风速剖面(Shimomai 等, 1996).

[6]中国华东地区雨天的年度变化(Jiang 等, 1997).

[7]英格兰中部地区的年平均温度(Baliunas 等, 1997).

作为本节所要讨论的一个特殊例子，考虑图 8 中的下图. 该图显示了关于一个特殊的铯原子钟的一系列测量. 该铯原子钟(代号 571)记录了每天的时间并且和美国海军天文台(这个数据将在 8.6 节和 9.7 节深入研究)拥有的时钟进行了比较，也就是测量每天的平均分形频率偏差. 关于这些偏离的精确定义在 Rutman (1978)中给出，但是这里它们也可以被视为信号 $x(\cdot)$在区间$\left[t-\frac{1}{2},\ t+\frac{1}{2}\right]$上的平均，这里的 t 以天计. 这些测量每天报告一次，因此可以对 t 在整数值上绘出$A(1,\ t)$.

如果对任何时间点 t 有 $A(1,\ t)=0$，那么 571 时钟在从 $t-\frac{1}{2}$开始到 $t+\frac{1}{2}$结束的一天中与主时钟相比既没有加快也没有减慢. 当 $A(1,\ t)<0$(如图 8 中所示的情况)时，571 时钟就减慢了，而从 $t-\frac{1}{2}$开始到 $t+\frac{1}{2}$结束时所减慢的时间数量与 $A(1,\ t)$ 成正比. 如果 $A(1,\ t)$ 总是等于一个负常数，则很容易纠正 571 时钟的减慢趋势(同理，如果我们知道新年的时候我们的表被设定到正确的时间，并且知道表每天减慢的精确的时间数，就可以通过查看当前表的时间和当天在距该年第一天的天数来纠正表的错误). 因此，一个好的反映 571 钟在天这样的尺度上保持正确时间的指标不是 $A(1,\ t)$，而是它在不同的时间周期上的变化. 如果将这种变化和两个平均的时间点联系在一起，就导致量

$$D\left(1,t-\frac{1}{2}\right)\equiv A(1,t)-A(1,t-1)=\int_{t-\frac{1}{2}}^{t+\frac{1}{2}}x(u)\,\mathrm{d}u-\int_{t-\frac{3}{2}}^{t-\frac{1}{2}}x(u)\,\mathrm{d}u,$$

或者等价地，

$$D(1,t)=A\left(1,t+\frac{1}{2}\right)-A\left(1,t-\frac{1}{2}\right)=\int_{t}^{t+1}x(u)\,\mathrm{d}u-\int_{t-1}^{t}x(u)\,\mathrm{d}u.$$

如果$|D(1,\ t)|$作为 t 的函数趋向于很大的值，那么 571 钟在天这样的尺度下就比较差. 如果$|D(1,\ t)|$接近于零，则时钟每天都很好. 更一般地，在尺度 λ 上，时钟的时间保持性能可以由下面的量来检验

$$D(\lambda,t)\equiv A\left(\lambda,t+\frac{\lambda}{2}\right)-A\left(\lambda,t-\frac{\lambda}{2}\right)=\frac{1}{\lambda}\int_{t}^{t+\lambda}x(u)\,\mathrm{d}u-\frac{1}{\lambda}\int_{t-\lambda}^{t}x(u)\,\mathrm{d}u. \quad (7)$$

对于其他过程，都有相似的结果：各种不同尺度下的均值的变化要比均值本身更重要(例如，当天与次日日平均气温的变化，海洋上年平均气温的变化等等). 例如，$D(1,\ t)$ 的图像能告诉我们某日与次日的日平均气温的变化速度. 同理，当尺度增大到年时，$D(\lambda,\ t)$ 的图像就告诉我们某年与次年的年平均气温的变化.

现在的目的是把小波和寻找均值的变化联系起来. 因为式(7)中的两个积分涉及两个相邻的不重合的区间，所以我们很容易把它们合写为整个实轴上的一个

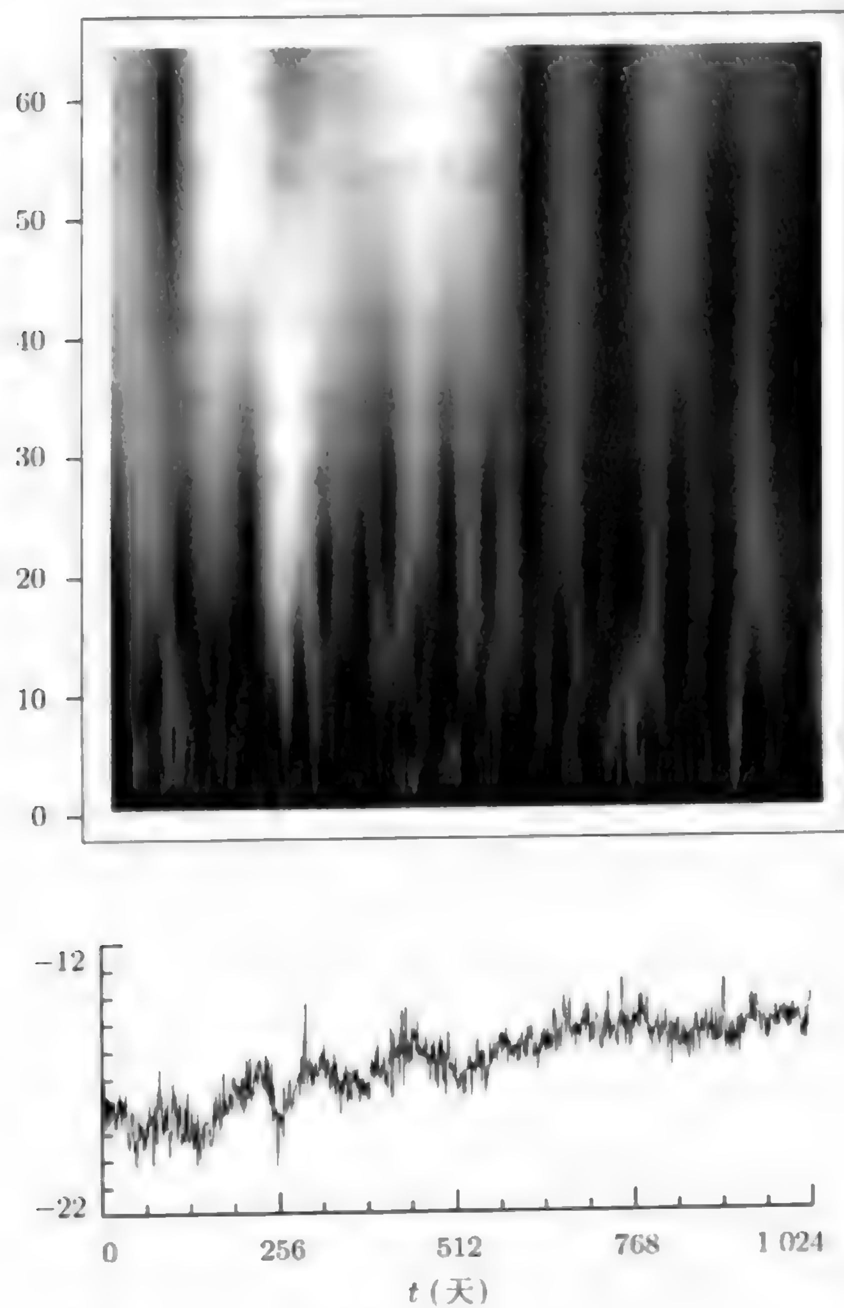


图 8 铯原子钟 571 每天的平均分形频率偏差(下图)和它的墨西哥帽连续小波变换. 分形频率偏差的记录误差是 10^{13} 数量级(偏差为 -15×10^{13} 秒时意味着 571 钟相对于美国海军天文台的时钟每天减慢十亿分之 129.6 秒). 上图是连续小波变换的图像，纵轴是尺度(从 1 天到 64 天)，横轴与下图是相同的. 连续小波变换是用灰度图像表示的，所以值越大所对应的点就越亮(图中黑色的区域意味着在对应的时间和尺度下 571 钟走得比较好)

积分：

$$D(\lambda,t)=\int_{-\infty}^{\infty}\widetilde{\varphi}_{\lambda,t}(u)x(u)\mathrm{d}u,$$

其中

$$\widetilde{\varphi}_{\lambda,t}(u)\equiv\begin{cases}-1/\lambda, & t-\lambda<u\leqslant t;\\ 1/\lambda, & t<u\leqslant t+\lambda;\\ 0, & 其他.\end{cases}$$

当 $\lambda=1$，$t=0$ 时取

$$\widetilde{\varphi}_{1,0}(u)\equiv\begin{cases}-1, & -1<u\leqslant 0;\\ 1, & 0<u\leqslant 1;\\ 0, & 其他.\end{cases}$$

如果将上式和式(2c)中定义的哈尔小波 $\psi^{(\mathrm{H})}(\cdot)$ 做比较，我们发现 $\widetilde{\varphi}_{1,0}(u)=\sqrt{2}\psi^{(\mathrm{H})}(u)$. 单位尺度下 $t=0$ 时的计算均值差分的格式等价于(差一个常数因子)信号 $x(\cdot)$ 和哈尔小波的积分. 从效果上来说，借助于积分

$$\int_{-\infty}^{\infty}\psi^{(\mathrm{H})}(u)x(u)\mathrm{d}u\equiv W^{(\mathrm{H})}(1,0),$$

哈尔小波提取了单位尺度下 $t=0$ 附近 $x(\cdot)$ 的两个均值的差的大小的信息.

通过调整哈尔小波，很容易提取关于单位尺度下任意时刻 t 处的变化的精确信息：只需要将 $\psi^{(\mathrm{H})}(\cdot)$ 平移即可. 相应地，定义

$$\psi_{1,t}^{(\mathrm{H})}(u)\equiv\psi^{(\mathrm{H})}(u-t),\quad 于是得到\ \psi_{1,t}^{(\mathrm{H})}(u)=\begin{cases}-1/\sqrt{2}, & t-1<u\leqslant t;\\ 1/\sqrt{2}, & t<u\leqslant t+1;\\ 0, & 其他.\end{cases}$$

(参见图 10 中上面一行.)因为 $\psi_{1,t}^{(\mathrm{H})}(\cdot)$ 是 $\psi^{(\mathrm{H})}(\cdot)$ 的平移，所以它显然满足小波的两条基本性质(积分为零和平方积分为 1). 对 $\psi_{1,t}^{(\mathrm{H})}(\cdot)$ 和信号 $x(\cdot)$ 的乘积进行积分得到

$$W^{(\mathrm{H})}(1,t)\equiv\int_{-\infty}^{\infty}\psi_{1,t}^{(\mathrm{H})}(u)x(u)\mathrm{d}u=\int_{t}^{t+1}x(u)\mathrm{d}u-\int_{t-1}^{t}x(u)\mathrm{d}u=D(1,t).$$

我们也可以提取其他关于尺度 λ 和时间 t 的类似的信息，考虑

$$\psi_{\lambda,t}^{(\mathrm{H})}(u)\equiv\frac{1}{\sqrt{\lambda}}\psi^{(\mathrm{H})}\left(\frac{u-t}{\lambda}\right)=\begin{cases}-\dfrac{1}{\sqrt{2\lambda}}, & t-\lambda<u\leqslant t;\\ \dfrac{1}{\sqrt{2\lambda}}, & t<u\leqslant t+\lambda;\\ 0, & 其他.\end{cases}$$

容易看出 $\psi_{\lambda,t}^{(\mathrm{H})}(\cdot)$ 满足式(2a)和(2b). 使用这个小波，得到

$$W^{(\mathrm{H})}(\lambda,t)\equiv\int_{-\infty}^{\infty}\psi_{\lambda,t}^{(\mathrm{H})}(u)x(u)\mathrm{d}u\propto D(\lambda,t).\tag{9}$$

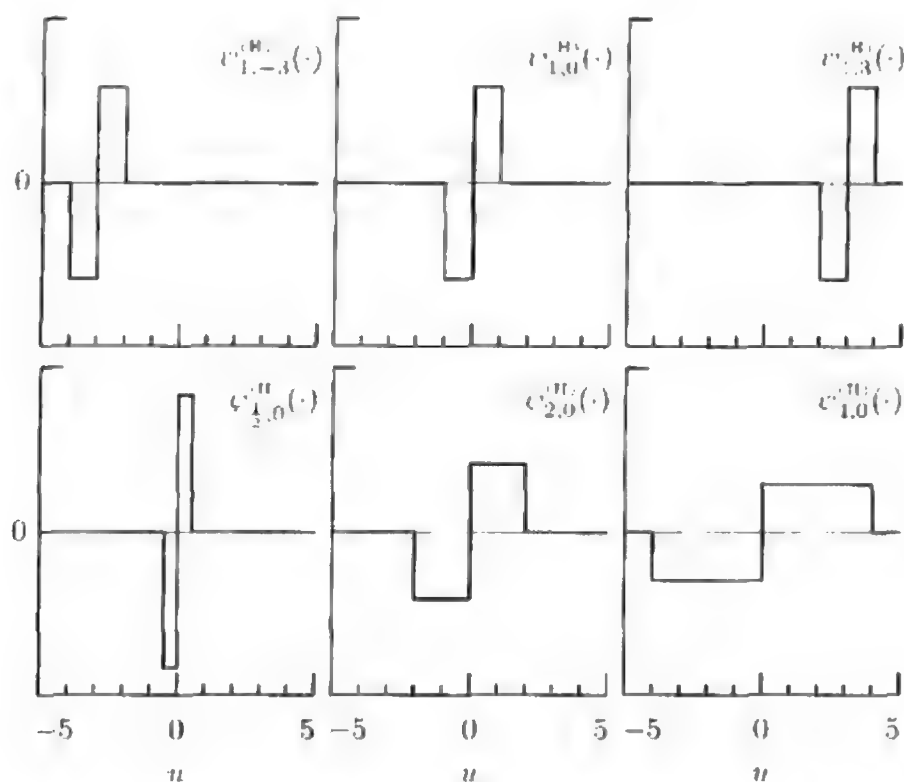


图 10 哈尔小波 $\psi^{(H)}(\cdot)$进行平移和变尺度后的例子. 图中所画的是函数 $\psi^{(H)}_{\lambda,t}(\cdot)$的图像，这个函数可以用来度量信号 $x(\cdot)$在 t 时刻在长度为 λ 的尺度下相邻两个区间均值的差. 图中，上面一行是尺度参数 λ 固定而时间参数 t 变化的情形；下面一行是时间参数 t 固定为零而尺度变化的情形

图 10 中最下面一行是 $\psi^{(H)}_{\lambda,0}(\cdot)$的三个例子.

通过改变 λ，可以绘制信号 $x(\cdot)$在不同尺度下从一个时段到下一时段(长为 λ)的均值变化的复杂图像. 变量形成的集合 $\{W^{(H)}(\lambda, t): \lambda>0, -\infty<t<\infty\}$ 就被作为已知信号 $x(\cdot)$的哈尔连续小波变换(CWT). 对这个变换的解释是：$W^{(H)}(\lambda, t)$ 与两个相邻的均值在 λ 尺度下的差成正比，第一个均值是从 t 时刻开始计算的，第二个均值是在 t 时刻结束的.

按照同样的方式，我们可以用其他的小波 $\psi(\cdot)$(例如图 3 中的 $\psi^{(fdG)}(\cdot)$或 $\psi^{(Mh)}(\cdot)$)并且基于它来构造连续小波变换，形式为

$$W(\lambda,t) \equiv \int_{-\infty}^{\infty} \psi_{\lambda,t}(u)x(u)\,\mathrm{d}u, \tag{10}$$

其中

$$\psi_{\lambda,t}(u) \equiv \frac{1}{\sqrt{\lambda}}\psi\left(\frac{u-t}{\lambda}\right).$$

对于小波 $\psi^{(fdG)}(\cdot)$和 $\psi^{(Mh)}(\cdot)$，我们可以按照哈尔连续小波变换类似的方式来解释 $W(\lambda, t)$ 告诉我们的关于 $x(\cdot)$的事情. 首先考虑图 3 中间的 $\psi^{(fdG)}(\cdot)$，我们可以断言：小波变换的结果就是两个相邻区间的加权平均的差分，而哈尔小波本质上就是观察两个相邻区间的均值的差分，这个均值是抽样的平均

$$\frac{1}{N}\sum_{j=0}^{N-1} x_j,$$

而 $\psi^{(\mathrm{fdG})}(\cdot)$ 小波涉及类似于

$$\frac{\sum_{j=0}^{N-1} w_j x_j}{\sum_{j=0}^{N-1} w_j}$$

的加权平均的差分(上面求差分的两个均值的权重是相互反转的，这是因为 $\psi^{(\mathrm{fdG})}(\cdot)$ 是反对称的). 另一方面，我们可以认为墨西哥帽小波 $\psi^{(\mathrm{Mh})}(\cdot)$ 生成了单位尺度上两个相邻加权平均的差分(参见下面的评论和扩展，是关于 $\psi^{(\mathrm{fdG})}(\cdot)$ 和 $\psi^{(\mathrm{Mh})}(\cdot)$ 小波可以认为是与单位尺度相关的讨论).

现在我们陈述一个有关连续小波变换的基本事实：连续小波变换的结果保留了原信号 $x(\cdot)$ 中的所有信息. 如果 $\psi(\cdot)$ 满足容许性条件(4a)，并且信号 $x(\cdot)$ 满足

$$\int_{-\infty}^{\infty} x^2(t)\,\mathrm{d}t < \infty,$$

那么，我们可以通过信号 $x(\cdot)$ 的连续小波变换来重构信号 $x(\cdot)$

$$x(t) = \frac{1}{C_\psi}\int_0^{\infty}\left[\int_{-\infty}^{\infty} W(\lambda,u)\frac{1}{\sqrt{\lambda}}\psi\left(\frac{t-u}{\lambda}\right)\mathrm{d}u\right]\frac{\mathrm{d}\lambda}{\lambda^2} \tag{11a}$$

(其中，C_ψ 由等式(4a)定义)；进而，我们还有

$$\int_{-\infty}^{\infty} x^2(t)\,\mathrm{d}t = \frac{1}{C_\psi}\int_0^{\infty}\left[\int_{-\infty}^{\infty} W^2(\lambda,t)\,\mathrm{d}t\right]\frac{\mathrm{d}\lambda}{\lambda^2} \tag{11b}$$

(Calderón, 1964; Grossmann and Morlet, 1984; Mallat, 1998, pp. 78-79). 上式的左边通常用来定义信号的 $x(\cdot)$ 的能量；上面的等式表明 $W^2(\lambda, t)/\lambda^2$ 在本质上定义了信号沿着尺度和时间分解的能量密度.

连续小波变换的基本应用是：将信号 $x(\cdot)$ 的某些特定的信息以一种新的方式呈现出来，使得我们可以得到关于 $x(\cdot)$ 的一些无法直接获取的信息(例如，$x(t)$ 随时间变化的图像). 因为我们可以通过信号的连续小波变换重构信号 $x(\cdot)$，所以信号和它的连续小波变换是同一个数学实体的两个不同的表现.

作为连续小波变换的例子，图 8 中的上图就是原子钟数据在尺度 λ 从 1 变化到 64 天时的图像. 在这个图像中，时间轴上短竖线标记出来的区间长度是 64 天，也就是连续小波变换的最大尺度. 所以，当尺度等于 64 的时候(也就是图的最上面一行)，在时间坐标 170 和 480 附近分别有大范围的亮点，表示时钟在区间范围(64 天)内在这些时刻点附近有特别大的变化. 当尺度等于 32 的时候，在时间等于 256 的附近有大范围的白色亮点，表示时钟在半个区间范围(32 天)内这些时刻点附近有特别大的变化. 这也可以通过时间序列图像(图 8 下)看出来. 在这个连续小波变换的灰度图像中，后一半图像要比前一半图像黑，这说明后一半时间里钟走得更稳定；在时刻 610 和 720 附近所有的尺度下图像都有一个黑带，这说明在这些时间里时钟走得特别稳定.

1.2 节的评论与扩展

[1]直观上看，图 3 中的哈尔小波 $\psi^{(H)}(\cdot)$用于度量单位尺度下的变化，因为正的部分

$$\psi^{(H)+}(u) \equiv \begin{cases} \psi^{(H)}(u) = 1/\sqrt{2}, & \psi^{(H)}(u) > 0, \text{即 } 0 < u \leqslant 1; \\ 0, & \text{其他} \end{cases}$$

类似于可给出的负的部分，两者在“物理”或视觉上看起来都是单位宽度．不过，另一个被大家接受的有关函数 $g(\cdot)$的宽度的度量是自相关宽度，定义为

$$\mathrm{width_a}\{g(\cdot)\} = \frac{\left|\int g(u)\mathrm{d}u\right|^2}{\int g^2(u)\mathrm{d}u}.$$

哈尔小波的正的部分有自相关宽度

$$\mathrm{width_a}\{\psi^{(H)+}(\cdot)\} = \frac{\left|\frac{1}{\sqrt{2}}\int_0^1 \mathrm{d}u\right|^2}{\frac{1}{2}\int_0^1 \mathrm{d}u} = 1,$$

同理，其负的部分也有相同的结果．对于式(3a)和(3b)定义的小波 $\psi^{(fdG)}(\cdot)$和 $\psi^{(Mh)}(\cdot)$，虽然它们的正的部分没有明显的“物理”的宽度，但是我们可以计算它们的自相关宽度．现在 $\psi^{(fdG)}(\cdot)$和 $\psi^{(Mh)}(\cdot)$的表达式都包含有一个调节参数 σ，为了使结果对 σ 的依赖关系明显，我们暂时将这两个小波记作 $\psi_\sigma(\cdot)$．于是可以写为

$$\psi_\sigma(u) = \frac{1}{\sqrt{\sigma}}\psi_1\left(\frac{u}{\sigma}\right).$$

令

$$\psi_\sigma^+(u) \equiv \begin{cases} \psi_\sigma(u), & \psi_\sigma(u) > 0; \\ 0, & \text{其他}. \end{cases}$$

那么，$\mathrm{width_a}\{\psi_\sigma^+(u)\} = \sigma \times \mathrm{width_a}\{\psi_1^+(u)\}$，其中 $\mathrm{width_a}\{\psi_1^+(u)\}$可以通过数值积分计算．令 $\sigma = 1/\mathrm{width_a}\{\psi_1^+(u)\}$，我们强制正的部分的宽度 $\mathrm{width_a}\{\psi_\sigma^+(u)\}$和哈尔小波一样，为单位 1．通过计算可得，对于 $\psi^{(fdG)}(\cdot)$，$\sigma \doteq 0.443\,11$；对 $\psi^{(Mh)}(\cdot)$，$\sigma \doteq 0.636\,28$，这些值正是我们在图 3 中绘制小波图像时所使用的参数．

[2]有关连续小波变换在自然科学中的有趣应用的入门读物有：Meyers 等(1993)，Kumar and Foufoula-Georgiou(1994)，Carmona 等(1998)以及 Torrence and Compo(1998)．

1.3　连续小波变换的延续：离散小波变换

任何基于连续小波变换分析信号都会出现信息的冗余．当绘制如图 8 中的图像时，连续小波变换本质上是一个探索型的数据分析工具，它可以帮助人眼从信号中寻找我们感兴趣的特征．为了从图像上获得更多的东西，我们就必然面临图像处理问题，因为连续小波变换在本质上是二维的．因为二维的连续小波变换只依赖于一维的信号，所以连续小波变换中显然有很多冗余．例如，在图 8 中当尺

度很大以后，两个相邻尺度范围内的连续小波变换的差异很小(比如说，比较一下尺度为 40 和 42 的两个细条)，在固定的大尺度范围内不同的时刻下变化很慢(如在尺度为 40 的范围内沿着细条从左到右移动时连续小波变换的变化)．因此，我们进一步利用二次抽样来获取一些关键性的特征．

本书的主要任务是讨论那些可以直接用于在离散时间点(也就是，$t=0, 1, \cdots, N-1$)上观察得到的时间序列的离散小波变换．正如我们将要在第 4 章看到的那样，离散小波变换完全可以不和连续小波变换联系而自己形成一套完整的体系(离散小波变换与连续小波变换的联系将在第 11 章中讨论)．而离散小波变换可以看作是希望以一种非常简洁的形式保留连续小波变换的主要特征的一种努力．从这种观点来看，离散小波变换可以看作是 $W(\lambda, t)$ 的巧妙的二次采样，而且我们只使用“二进”尺度(即尺度 λ 的形式是 2^{j-1}，$j=1, 2, 3, \cdots$)，在特定的尺度 2^{j-1} 内，我们以步长为 2^j 来离散时间．

作为一个例子，图 14 中左边部分的图像显示了时钟数据的离散哈尔小波变换的结果．如图中左半部分最下图(也就是图 8 中的下图)所示，数据中有 $N=1\,024$ 个值．这些数据的离散小波变换也有 $N=1\,024$ 个数据，称为“离散小波变换系数”．关于时钟数据的这些系数可以整理成 8 个序列，它们都绘制在图中左半部分并排成一列．这些序列中的 7 个序列称为“小波系数”——它们可以看作是在“尺度 1”往上到“尺度 64”各图中与零的偏差．对于尺度 2^{j-1}，有 $N_j\equiv N/2^j$ 个小波系数，对应这些系数的时间点可以取作 $(2n+1)2^{j-1}-1/2$，$n=0, 1, \cdots, N_j-1$ (对于每一个尺度，小波系数的个数和时间间隔的乘积依然是 N，也就是原始数据的长度)．通过合适的规范过程，尺度为 2^{j-1} 的第 n 个小波系数可以认为是 $W^{(H)}\left(2^{j-1}, (2n+1)2^{j-1}-\frac{1}{2}\right)$的一个逼近，即式(9)的哈尔连续小波变换．

上述这 7 个尺度的小波系数总计是 1 016 个离散小波变换系数．剩余的 8 个系数就是所谓的“尺度系数”，在图 14 左半部分的最上一个图中用直线连在一起(与尺度为 64 的小波系数在同样的时间坐标下)．尺度系数与原始数据在尺度为 128 时的均值成正比(与小波系数不同，小波系数与两个相邻均值的差分成正比)．所以，尺度系数反映的是长期的变化，这就是它们为什么显示出与原始数据类似的向上趋势．

对图 14 中小波系数的解释与对 $W(\lambda, t)$ 的解释是十分相似的．比如说，在尺度为 64 的时候，最大的系数是标号为 $n=1$ 的系数．与之相联系的时间是 $(2n+1)\cdot 64-1/2=191.5$．这个系数反映了在时间 191.5 之前和之后的尺度为 64 的均值之间的区别．这个差值很大的事实表明：当尺度为 64 时，与其他时间段相比，时间 $t=191.5$ 附近的钟的行为比较差．回忆一下图 8 中墨西哥帽连续小波变换，我们发现连续小波变换和小波系数有相似的模式．比如说，两个图都显示出在尺度为 16 和 32 的时候钟的行为似乎是随着时间的转移而越来越好；另一方面，在最小的尺度下，钟的行为与时间变化是相当一致的(参见 9.6 节中关于统计检验的讨论，我们将客观地判断小波系数在特定的尺度下

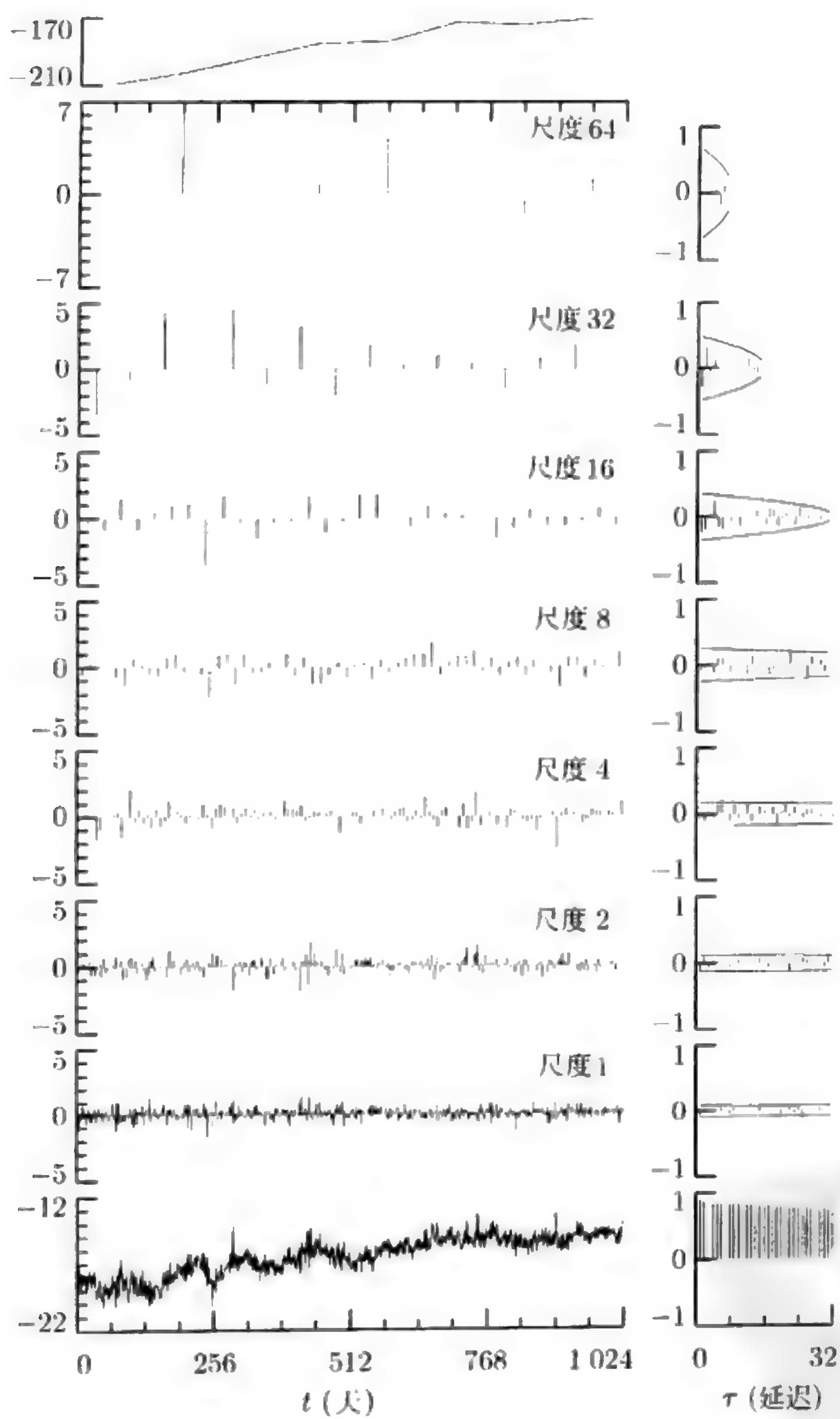


图 14 时钟 571 的哈尔离散小波变换系数和样本自相关序列

随着时间的改变是否有相同的变化).

正如信号 $x(\cdot)$可以用它自身的连续小波变换重构一样(参见式(11a)),时钟的数据也可以利用它的离散小波变换系数进行重构(详见第 4 章中的讨论).因此,连续小波变换在二进尺度上的二次抽样似乎是进行了大幅度的约简,但是时间序列和它的离散小波变换是同一个数学实体的两种不同的表示形式.于是,我们可以断言,在处理抽样的时间序列的时候(最近这种收集时间序列的方式是很流行的),将连续小波变换变为离散小波变换不会丢失任何信息.事实上,尤其是在科学领域,标准的二进尺度在一般情况下就足够特征化一个物理过程.在科学领域中很长时间里人们都自然地根据多个物理尺度(例如,大气科学、海洋学、天文学等)来描述演化过程.(如果用离散小波变换给出的连续小波变换的标准二进尺度概括对于某些问题并不恰当,则还有很多修改后的变换——如离散小波包变换——它可以提供更合适的分析.与之相关的思想将在第 6 章讨论.)

作为连续小波变换的缩影,离散小波变换保留了在时间序列中按尺度分解能量的能力.如果令 X_t、$W_{j,n}$和 V_n 分别表示第 t 个数据值、尺度 2^{j-1}的第 n 个小波系数和第 n 个尺度系数,则有

$$\sum_{t=0}^{1\,023} X_t^2 = \sum_{j=1}^{7}\sum_{n=0}^{N_j-1} W_{j,n}^2 + \sum_{n=0}^{7} V_n^2,$$

这个结果是式(11b)中能量分解的离散形式.上式中,双重求和中的第一次求和是时间序列在尺度 2^{j-1}中的能量求和,各个尺度的能量沿着尺度的变化在图 16 中用直线连接并显示了出来.时钟分析师从这幅图上可以看出 571 时钟在 8 天到 16 天这样的尺度上走得最好(这将会影响在实际中,比如说在测地学中,如何使用这个钟).

正如将要在第 8 章中讨论的,能量分解与小波方差(wavelet variance)的概念是密切相关的,在小波方差中,我们将过程中的变量按照二进尺度分割.小波方差是基于傅里叶变换的能量谱的一个简洁的替换物,作为基于尺度的分析方法,小波方差比基于频率的频谱更容易解释.在统计模型的帮助下,我们可以使用小波方差的各种估计方法估计抽样变异(sampling variability)效应.例如,这样做可以建立假设为真的小波方差的置信区间(这允许我们像图 16中一样估计图形中的抽样变异效应).另外,正如第 8 章中所演示的例子一样,我们可以利用小波系数的局部化性质去分析过程随时间的演变.这种方法允许我们研究与时钟数据不同的、具有尺度特征的时间序列,它们随着时间的改变是不相似的.

现在,让我们考虑离散小波的另一个重要的方面,如图 14 中右半列所示.底部最大的那个图显示了时间序列的样本自相关序列的第一部分,样本自相关序

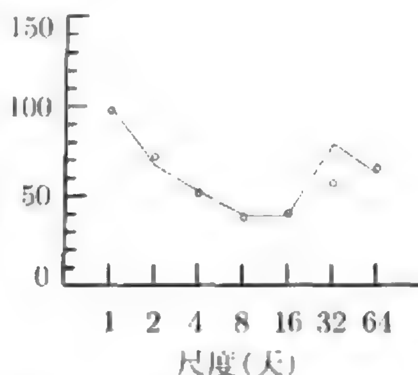


图 16　基于哈尔离散小波变换小波系数(曲线)和极大覆盖哈尔离散小波变换小波系数(小圆圈)的时钟 571 的能量分析

列如下定义：

$$\hat{\rho}_{X,\tau} \equiv \frac{\sum_{t=0}^{N-\tau-1}(X_t-\overline{X})(X_{t+\tau}-\overline{X})}{\sum_{t=0}^{N-1}(X_t-\overline{X})^2}, \quad \tau=1,2,\cdots,32,$$

其中，$\overline{X}$ 是时钟数据的样本均值. 统计量 $\hat{\rho}_{X,\tau}$ 是时间序列与它的 τ 个单位的平移之间的线性关联（association）的度量. 这个统计量的范围是 $-1<\hat{\rho}_{X,\tau}<1$. 当 $\hat{\rho}_{X,\tau}$ 接近于 1 的时候，X_t 和 $X_{t+\tau}$ 的值就非常近(与时间序列的总变异性相比). 时钟数据的 $\hat{\rho}_{X,\tau}$ 图像显示它有高度的相关性，即使是间隔为 32 天的偏差也是如此(事实上，当 $\tau=32$ 的时候样本自相关序列的值会慢慢地趋向于零. 这种缓慢衰减是一种典型的“长记忆过程”——在 7.6 节中我们将给出这种过程中的基本概念的综述).

对给定的尺度，小波系数自然地形成了一个特有的时间序列，所以感兴趣的是研究这些系数的样本自相关序列. 在图 14 右半部分上面的七个图像显示了部分自相关序列与零的偏差的图像(由于最大的三个尺度的自相关序列的数量比较少，所以它们的图像是全部自相关序列的图像). 这里，我们可以发现，样本自相关函数的数值是非常小的. 在某个温和的合理假设下，统计理论告诉我们，如果小波系数的自相关系数是 0，那么尺度为 2^{j-1} 的小波系数的样本自相关序列中有 95% 的数会落在 $\pm 2\sqrt{N_j-\tau}/N_j$ 之间(详情参见 Fuller, 1996，推论 6.3.6.2). 这些界限在图中也和样本自相关序列一起绘制了出来. 我们看到，当考虑样本变异性的时候，很难做出小波系数的自相关非零的判断(对不同尺度的小波系数的相关性我们也有类似的结论). 因此，尽管时钟原始数据显示出了高度的自相关性，但是小波系数几乎是不相关的.

时间序列统计分析的中心目的是寻找将相关序列表示成不相关序列的组合的方法. 图 14 中，自相关序列的图像告诉我们离散小波变换能有效地将相关序列(甚至是高度相关的序列)去相关，结合我们可以利用离散小波变换的系数精确地重构原来的时间序列——这就是离散小波变换是时间序列分析的一个很有用的工具的最根

本原因. 在第 8 章(小波方差)、第 9 章(长记忆过程的分解和合成)和第 10 章(信号估计)中讨论的统计方法都依赖于离散小波的精确重构和去相关性. 在小波领域中，很多发展都可以归结为这两个重要的性质和每一个小波系数都只依赖于部分时间序列的事实(这就给我们有效地处理其统计特征随着时间的变化而演化的时间序列提供了一种可能). (正如第 4 章将要讨论的那样，另一个"卖点"是完全可以用离散小波变换计算出来的. 借助于金字塔算法，计算离散小波变换有一种十分巧妙的方法. 从计算速度的角度上来说，这种算法要比快速傅里叶变换还要快，而快速傅里叶变换使得离散傅里叶变换从 20 世纪 60 年代开始广为使用.)

最后，我们注意到，虽然离散小波变换的二次抽样格式有效地将二维连续小波变换降为一维运算，但是有很多的时间序列可以利用在时间上相对较细致的抽样更加有效地进行处理. 对这些序列，在二次抽样的时候保留一定的连续小波变换的冗余是很有好处的. 在第 5 章中我们将讨论一种离散小波变换的修改形式，称为极大覆盖离散小波变换(maximal overlap DWT，MODWT). 和离散小波变换一样，极大覆盖离散小波变换可以认为是连续小波变换在二进尺度上的二次抽样，但是与离散小波变换不同的是，我们处理的是所有时间轴上的信号，而不仅仅是 2^j 间隔的采样点上的值. 保留所有时间点上的值可以得到更好的连续小波变换的适当简化形式，因为它可以消除离散小波变换对连续小波变换在时间上进行二次抽样的过程中造成的一定的"对准"效应(alignment artifact).

作为一个例子，图 18 中显示了时钟数据的哈尔极大覆盖离散小波变换的小波系数(中间七个图像)和尺度系数(最上面一个图像). 与离散小波变换相比，极大覆盖离散小波变换的每一个尺度都有 1 024 个小波系数，它的尺度系数也有 1 024 个. 通过适当规范化，尺度 2^{j-1} 的第 t 个极大覆盖离散小波变换的小波系数可以认为是式(9)中 $W^{(H)}(2^{j-1}, t)$ 的一个逼近. 如果选择一个好的方法进行二次抽样和重新规范化，离散小波变换的所有小波系数可以从极大覆盖离散小波变换的系数中得到. 我们很容易利用极大覆盖小波变换的系数去重构时钟的原始数据. 如果我们用$\widetilde{W}_{j,t}$和$\widetilde{V}_t$分别表示尺度为 2^{j-1} 的第 t 个极大覆盖小波系数和第 t 个尺度系数，则有下面的能量分解：

$$\sum_{t=0}^{1\,023} X_t^2 = \sum_{j=1}^{7} \sum_{t=0}^{1\,023} \widetilde{W}_{j,t}^2 + \sum_{t=0}^{1\,023} \widetilde{V}_t^2.$$

正如在离散小波变换中的那样，双重求和中的第一次求和可以认为是在尺度 2^{j-1} 下能量贡献的求和. 在图 16 中我们用小圆圈画出了这些和. 这个分解将在第 8 章中有很好的应用，在那里我们将用它建立基于极大覆盖离散小波变换的基于小波变量的估计量. 得到的估计量在统计意义下要比对应的基于离散小波变换的估计量要更加有效，这是极大覆盖离散小波变的最重要的优点(但是，这个变换的计算速度要比离散小波变换慢. 事实上，算法的复杂性和快速傅里叶变换相同).

总之，尽管离散小波变换可以看作是连续小波变换的二次采样，但是实际上可以认为，它是独立的并且是一个重要而又实用的进行时间序列分析的工具. 离散小波变换是一个有效的分析工具的基本理由如下：

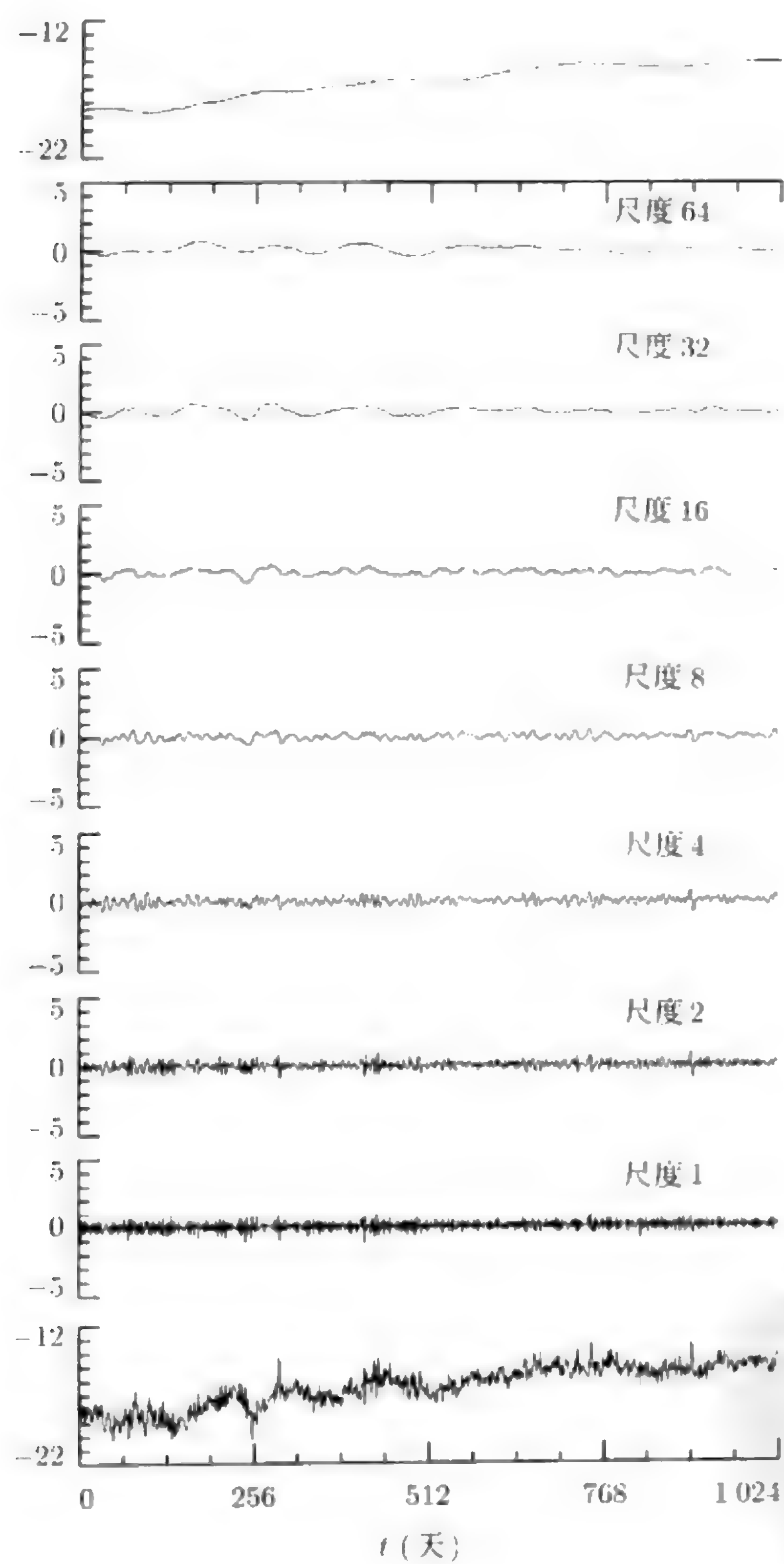


图 18 时钟 571 的哈尔极大覆盖离散小波变换系数

[1]离散小波变换将时间序列重新表示为与特定的时间和特定的二进尺度2^{j-1}相关联的系数形式. 由于可以从它的离散小波变换的系数重构时间序列，所以这些系数就等价于原来的序列.

[2]离散小波变换允许我们将时间序列的能量分割为与不同的时间和尺度相关联的片段. 能量分解和著名的统计中的方差分析(ANOVA)技术很相似，于是利用离散小波变换就可以得到基于尺度的方差分析，这与利用能量谱而建立的基于频率的方差分析是十分相似的.

[3]离散小波变换能有效地将物理应用中常见的大量的时间序列去相关. 这个性质是将在第8、9和10章中讨论的基于小波的统计方法的关键.

[4]离散小波变换可以通过一种比快速傅里叶变换更快的算法进行计算.

对离散小波变换的一种修改产生了极大覆盖离散小波变换. 这种变换在时间方向上引入了一些冗余，从而可以非常好地分析一些特定的时间序列. 上面所有这些发展和与之相关的内容——从最基本的开始——就是本书第4章到第10章的主要内容.

1.3 节的评论与扩展

[1]为了保证内容的完整性，需要对我们所分析的时钟数据进行一些详细的介绍(其中一些数据与本书后面的一些内容有关). 时钟数据事实上有1 025个数值，但是在本章的分析中只用了前1 024个. 正如将在第4.7节和第5章中所说的那样，本章所说的变换实际上是局部的$J_0=7$层离散小波变换和局部的极大覆盖离散小波变换. 在计算局部的极大覆盖离散小波变换的时侯，我们要使用反射边界条件(参见式(140)附近的讨论). 由于第8章中的原因，我们可以假设在这个过程中每一个尺度的小波系数的均值都是零，这样就可以通过公式

$$\frac{\sum_{t=0}^{N_j-\tau-1} \mathbf{W}_{j,t}\mathbf{W}_{j,t+\tau}}{\sum_{t=0}^{N_j-1}\mathbf{W}_{j,t}^2}$$

来估计这些系数的自相关系数的第τ个单位.

第 2 章

傅里叶理论和滤波器回顾

2.0 引言

在接下来的几章中，我们将对取自序列傅里叶变换和函数傅里叶变换理论中的一些基本结果做具体的应用，并且我们将会发现滤波器在小波应用方面起到中心的作用. 本章计划作为一个自我包容的，导出来自傅里叶理论和滤波理论的关键结果. 我们所选择的材料有意识地局限在我们后面将要用到的内容. 对于这里采用的符号和变换的具体讨论，详见 Percival and Walden(1993). 我们也推荐 Briggs and Henson(1995)以及 Hamming(1989)作为进一步阅读的补充材料.

熟知傅里叶分析和滤波器的读者可以很快地浏览本章，熟悉一下我们的符号和约定即可. 我们鼓励其他读者认真地学习本章并且尽可能多地做书中嵌入的练习(答案见附录). 这对于理解将在 2.6 节讲解的周期滤波器的概念特别重要，因为我们将在第 4 章和第 5 章反复地使用这一概念.

2.1 复变量与复指数

序列和函数的傅里叶理论的最优美的形式涉及复变量的应用，因此我们回顾一些关键的概念(详细的处理，可参见 Brown and Churchill，1995). 令 $i\equiv\sqrt{-1}$，因此 $i^2=-1$(在本书中，我们用“$\equiv$”表示“定义相等”). 如果 x 和 y 是实值变量，那么 $z\equiv x+iy$ 定义了一个复值变量，它的实部和虚部分别是 x 和 y. 本质上，z 是一个二维的变量，采用虚数 i，可以分别描述它的分量. 我们注意到下面的关键性质和定义(x，x_1，x_2，y，y_1，y_2 都假设为实值变量).

[1]两个复值变量 $z_1\equiv x_1+iy_1$ 和 $z_2\equiv x_2+iy_2$ 相等当且仅当它们的实部和虚部分别相等，即 $z_1=z_2$ 当且仅当 $x_1=x_2$ 且 $y_1=y_2$.

[2]$z_1+z_2=x_1+x_2+i(y_1+y_2)$且 $z_1z_2=x_1x_2-y_1y_2+i(x_1y_2+x_2y_1)$.

[3]$z=x+\mathrm{i}y$ 的复共轭是 $z^*\equiv x-\mathrm{i}y$. 注意到$(z_1+z_2)^*=z_1^*+z_2^*$; $(z_1z_2)^*=z_1^*z_2^*$; $x^*=x$(即实变量的复共轭是它自身).

[4]$z=x+\mathrm{i}y$ 的模(或绝对值, 或大小)是非负数 $|z|\equiv(x^2+y^2)^{1/2}$; 平方模是 $|z|^2=x^2+y^2$. 注意到 $zz^*=|z|^2$, $|z^*|=|z|$, $|z|^2\neq z^2$ 除非 $y=0$(即除非 z 可以看成是实值)和 $|z_1z_2|=|z_1||z_2|$. 如果我们在平面上画出点(x, y), 那么 $|z|$ 是该点与原点(0, 0)之间的距离.

[5]复指数 $\mathrm{e}^{\mathrm{i}x}$ 定义为一个复变量, 它的实部和虚部是 $\cos x$ 和 $\sin x$:

$$\mathrm{e}^{\mathrm{i}x}\equiv\cos x+\mathrm{i}\sin x.$$

上式就是著名的欧拉(Euler)公式. 注意到 $|\mathrm{e}^{\mathrm{i}x}|^2=\cos^2x+\sin^2x=1$ 对于所有的 x 成立. 复指数遵循与普通指数相似的计算规则, 包括:

(a)$\mathrm{e}^{\mathrm{i}(x+y)}=\mathrm{e}^{\mathrm{i}x}\mathrm{e}^{\mathrm{i}y}$.

(b)$(\mathrm{e}^{\mathrm{i}x})^n=\mathrm{e}^{\mathrm{i}nx}$, 这里 n 是任意的整数(这就是著名的棣莫弗定理).

(c)$\dfrac{\mathrm{d}\mathrm{e}^{\mathrm{i}x}}{\mathrm{d}x}=\mathrm{i}\mathrm{e}^{\mathrm{i}x}$, $\int\mathrm{e}^{\mathrm{i}x}\mathrm{d}x=\mathrm{e}^{\mathrm{i}x}/\mathrm{i}$.

[6]正弦和余弦函数可以用复指数的形式表示如下:

$$\cos x=\frac{\mathrm{e}^{\mathrm{i}x}+\mathrm{e}^{-\mathrm{i}x}}{2},\quad \sin x=\frac{\mathrm{e}^{\mathrm{i}x}-\mathrm{e}^{-\mathrm{i}x}}{2\mathrm{i}}.$$

[7]欧拉公式暗示了 $\mathrm{i}=\mathrm{e}^{\mathrm{i}\pi/2}$, $-\mathrm{i}=\mathrm{e}^{-\mathrm{i}\pi/2}$ 和 $-1=\mathrm{e}^{\pm\mathrm{i}\pi}$. 这些看起来不起眼的事实在化简不同的表达式时是很有用的.

[8]除了笛卡儿坐标表示 $z=x+\mathrm{i}y$, 任何复值变量都有一个极坐标表示, 记为 $z=|z|\mathrm{e}^{\mathrm{i}\theta}$, 这里和前面一样, $|z|\equiv(x^2+y^2)^{1/2}$, θ 叫做 z 的辐角(通常写为 $\arg z=\theta$ 且仅对 $|z|>0$ 有意义). 如果我们采用约定 $-\pi<\theta\leqslant\pi$, 并且在图上从点(x, y)到原点(0, 0)画一条直线, 那么 θ 表示该直线和 x 轴正向的两个夹角中的较小的一个(当 $y>0$ 时我们有 $\theta>0$, 当 $y<0$ 时$\theta<0$). 我们有 $\tan\theta=y/x$, 因此 θ 可以通过一个反正切函数计算, 要注意 x 和 y 的符号(即点(x, y)所在的象限). z 用实值变量 x, y 表示的笛卡儿坐标形式与用非负变量 $|z|$ 和满足 $-\pi<\theta\leqslant\pi$ 的辐角 θ 表示的极坐标形式是完全等价的.

[9]对于 $z=x+\mathrm{i}y$, $\Re(z)\equiv x$ 和$\Im(z)\equiv y$ 分别表示 z 的实部和虚部的函数.

2.2 无限序列的傅里叶变换

令$\{a_t: t=\cdots, -1, 0, 1, \cdots\}$表示一个实值或复值变量的无限序列, 满足 $\sum_{t=-\infty}^{\infty}|a_t|^2<\infty$(当 t 所假定的值可以从上下文清楚地知道时我们简记为$\{a_t\}$). 平方模的和为一个有限数的假设是一个充分但不是必要的条件, 使得我们在本节中所处理的所有量都是有定义的. $\{a_t\}$的离散傅里叶变换(DFT)是一个由下式定义的复

值函数：

$$A(f) \equiv \sum_{t=-\infty}^{\infty} a_t \mathrm{e}^{-\mathrm{i}2\pi ft},$$

其中$-\infty < f < \infty$是一个称作频率的实值变量. 回想一下，对于$f>0$，由$\cos(2\pi fu)$和$\sin(2\pi fu)$，$-\infty<u<\infty$定义的关于u的函数叫做关于频率f的正弦曲线，在这里频率有精确的定义，即当连续变量u从0变化到1时正弦的循环次数或百分数. 由欧拉公式可得

$$\mathrm{e}^{-\mathrm{i}2\pi fu} = \cos(2\pi fu) - \mathrm{i}\sin(2\pi fu),$$

自然$\mathrm{e}^{-\mathrm{i}2\pi fu}$定义的关于$u$的函数称为一个关于频率$f$的复指数函数，这里$|f|$表示当$u$从0变到1时函数的实部和虚部循环了多少次. 注意到在$A(\cdot)$的定义中，我们允许f是负的：任何负频率应该仅仅视为一种数学上的构造，事实上是为了当需要物理解释时可以和某些正频率相对应(例如，考虑练习[2.1]在复解调中的关于负频率的解释). 当在离散傅里叶变换的定义中使用$\mathrm{e}^{-\mathrm{i}2\pi ft}$时会扰乱最初的意思，读者可以放心，所涉及的正弦和余弦函数都是实的.

函数$A(\cdot)$有时也叫做$\{a_t\}$的傅里叶变换. 直观地，$A(f)$可以看成是一个尝试，即序列$\{a_t\}$和$\{\mathrm{e}^{-\mathrm{i}2\pi ft}\}$相互匹配得怎么样. 如果$|A(f)|$很大(很小)，那么匹配是好(差)的.

下面的两个练习记录了离散傅里叶变换的重要性质.

练习[22a] 证明$A(\cdot)$是单位周期的，即对于任意的整数j，我们有$A(f+j)=A(f)$. ◁

由于它的周期性，我们只需要考虑$A(\cdot)$在任一单位长度区间上的情况. 一个简便的选择就是闭区间$[-1/2, 1/2]$；即$|f|\leqslant 1/2$. 频率1/2通常叫做奈奎斯特(Nyquist)频率.

练习[22b] 证明：如果$\{a_t\}$是实值的，那么$A(-f)=A^*(f)$. ◁

由于$|A^*(f)|=|A(f)|$，由此得出实值序列的傅里叶变换的模是一个偶函数，即$|A(-f)|=|A(f)|$.

傅里叶变换的逆变换由下面的练习给出.

练习[22c] 证明：

$$\int_{-1/2}^{1/2} A(f)\mathrm{e}^{\mathrm{i}2\pi ft}\,\mathrm{d}f = a_t, \quad t=\cdots,-1,0,1,\cdots. \quad ◁$$

上面的公式说明我们可以从傅里叶变换$A(\cdot)$重构序列$\{a_t\}$. 所以，在很强的意义上说，对于相同的数学实体，$\{a_t\}$和它的离散傅里叶变换是两种表示形式. 大致来说，离散傅里叶变换的逆变换表明可以用正弦曲线的线性组合来重新表示$\{a_t\}$. 在重构$\{a_t\}$的过程中，$A(f)$的$|A(f)|$越大，正弦曲线f的频率就

越重要. $\{a_t\}$和$A(\cdot)$之间的傅里叶变换关系如下所示:

$$\{a_t\} \leftrightarrow A(\cdot),$$

也就是

$$A(f) = \sum_{t=-\infty}^{\infty} a_t e^{-i2\pi ft}$$

和

$$a_t = \int_{-1/2}^{1/2} A(f) e^{i2\pi ft} \, df,$$

这是用无穷和与积分定义的. 序列$\{a_t\}$和函数$A(\cdot)$称为傅里叶变换对. 这样偶尔可以很方便地使用记号$\{a_t\} \leftrightarrow A(f)$, 即使这种方法使我们混淆了函数$A(\cdot)$和它在$f$点的值$A(f)$的区别.

我们在以下练习中注意到一个特别重要的结果.

练习[23a] 假设有$\{a_t\} \leftrightarrow A(\cdot)$及$\{b_t\} \leftrightarrow B(\cdot)$, 其中$\sum_{t=-\infty}^{\infty} |a_t|^2 < \infty$并且$\sum_{t=-\infty}^{\infty} |b_t|^2 < \infty$. 导出这"两个序列"的帕塞瓦尔(Parseval)定理变形, 即

$$\sum_{t=-\infty}^{\infty} a_t b_t^* = \int_{-1/2}^{1/2} A(f) B^*(f) \, df.$$ ◁

上述结论的直接推论是"一个序列"的帕塞瓦尔(Parseval)定理

$$\sum_{t=-\infty}^{\infty} |a_t|^2 = \int_{-1/2}^{1/2} |A(f)|^2 \, df.$$

帕塞瓦尔定理指出一个序列的关键性质(平方和)可以通过它的傅里叶变换继续保持.

我们有机会来讨论以下的结果, 这个结果可以把一个序列的偶指标变量的离散傅里叶变换和整个序列的离散傅里叶变换联系起来.

练习[23b] 假设$\{a_t: t=\cdots, -1, 0, 1, \cdots\} \leftrightarrow A(\cdot)$, 证明一个无限序列$\{a_{2n}: n=\cdots, -1, 0, 1, \cdots\}$的离散傅里叶变换是由$\frac{1}{2}\left[A\left(\frac{f}{2}\right) + A\left(\frac{f}{2} + \frac{1}{2}\right)\right]$定义的函数. ◁

最后, 假设$\{a_t: t=0, 1, \cdots, N-1\}$是变量$N$的一个有限序列, 并且假定通过定义当$t \leqslant -1$或$t \geqslant N$时$a_t = 0$把它扩展为一个无限序列时. 这个序列的离散傅里叶变换由下式给出:

$$A(f) \equiv \sum_{t=-\infty}^{\infty} a_t e^{-i2\pi ft} = \sum_{t=0}^{N-1} a_t e^{-i2\pi ft}.$$

因此使用记号

$$\{a_t: t = 0, \cdots, N-1\} \leftrightarrow A(\cdot)$$

意味着

$$A(f) \equiv \sum_{t=0}^{N-1} a_t e^{-i2\pi ft}, \quad |f| \leqslant \frac{1}{2};$$

即我们默认为对$0, \cdots, N-1$以外的指标t, $a_t = 0$. 如果a_t的指标的范围由上下文可以清楚地得到, 就可以使用简化符号$\{a_t\} \leftrightarrow A(\cdot)$.

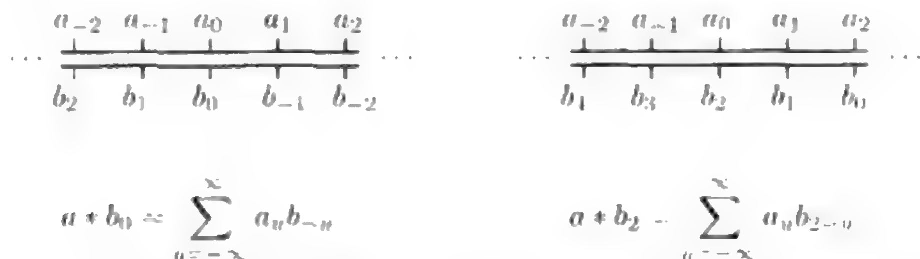


图 24 无限序列$\{a_t\}$和$\{b_t\}$的卷积如图所示. 左图中有两条线. 上面一条是无限序列$\{a_t\}$的元素均布顺序排列. 下面一条线也是均布排列的, 不过是无限序列$\{b_t\}$的元素的逆序排列, 也就是说, 无限序列$\{a_t\}$和$\{b_t\}$的卷积是以$\{b_{-t}\}$的首元素 $a * b_0$ 开始然后把数列$\{a_t\}$和$\{b_t\}$对应的元素依次相乘然后相加得到的. 总的来说, 第 t 个元素 $a * b_t$ 是由下面一行的元素由 t 分割以后转变为右边的形式以后通过相似的方法得到的——举例来说, 右图表示以第 2 个元素 $a * b_2$ 产生的卷积

2.3 无限序列的卷积/滤波

假定$\{a_t\}$和$\{b_t\}$是两个实数或复数变量组成的无限序列, 满足条件$\sum_{t=-\infty}^{\infty} |a_t|^2 < \infty$以及$\sum_{t=-\infty}^{\infty} |b_t|^2 < \infty$. 定义$\{a_t\}$和$\{b_t\}$的卷积为如下的无限序列, 其第 t 个元素是

$$a * b_t \equiv \sum_{u=-\infty}^{\infty} a_u b_{t-u}, \quad t = \cdots, -1, 0, 1, \cdots \tag{24}$$

(为简化我们的讨论, 假定贯穿本节的所有卷积$\{a * b_t\}$满足 $\sum_{t=-\infty}^{\infty} |a * b_t|^2 < \infty$, 但是重要的是要注意到这在一般情况下并不需要都为真). 注意到第二个序列$\{b_t\}$在做元素和元素的乘法之前是反转的并且关于第一个序列做了平移, 这种过程在图 24 中已经列举出来了.

下述练习或许会指出卷积中最重要的事实.

练习[24] 如果$\{a_t\} \leftrightarrow A(\cdot)$且$\{b_t\} \leftrightarrow B(\cdot)$, 证明卷积$\{a * b_t\}$的离散傅里叶变换可以由 $A(f)B(f)$定义的函数表示, 即

$$\sum_{t=-\infty}^{\infty} a * b_t e^{-i2\pi ft} = A(f)B(f), \quad |f| \leqslant \frac{1}{2}. \quad \triangleleft$$

对于这个结果的一种解释就是"时域范围内的卷积等同于频域范围内的相乘". 使用我们所定义的傅里叶变换的记号, 可以写为

$$\{a * b_t\} \leftrightarrow A(\cdot)B(\cdot),$$

其中 $A(\cdot)B(\cdot)$表示函数在 f 的值是 $A(f)B(f)$.

和卷积密切相关的一个概念就是复互相关, 它由如下无限序列来定义:

$$a^* \star b_t \equiv \sum_{u=-\infty}^{\infty} a_u^* b_{u+t}, \quad t = \cdots, -1, 0, 1, \cdots.$$

复互相关是$\{a_t^*\}$与$\{b_t\}$关于时间的卷积. 练习[2.2]指出

$$\{a^* \star b_t\} \leftrightarrow A^*(\cdot)B(\cdot);$$

即

$$\sum_{t=-\infty}^{\infty} a^* \star b_t e^{-i2\pi ft} = A^*(f)B(f). \tag{25a}$$

令上面的$b_t = a_t$, 可以得出$\{a_t\}$的自相关函数, 它是如下定义的无限序列:

$$a^* \star a_t \equiv \sum_{u=-\infty}^{\infty} a_u^* a_{u+t}, \quad t = \cdots, -1, 0, 1, \cdots. \tag{25b}$$

练习[2.2]的一个直接推论就是

$$\{a^* \star a_t\} \leftrightarrow A^*(\cdot)A(\cdot) = |A(\cdot)|^2, \tag{25c}$$

其中$|A(\cdot)|^2$表示在f的值是$|A(f)|^2$的函数. 因此一个自相关函数的离散傅里叶变换不仅是实值的, 而且也必然是非负的.

假如我们能够正确表示卷积的组成分量, 就能得到工程上的滤波器的概念(从技术上来说, 我们所说的滤波器是线性的时间不变的滤波器; 例如, 详见Percival and Walden, 1993). 因此, 如果我们把式(24)中的$\{a_t\}$作为一个滤波器并把$\{b_t\}$作为待滤波的数列, 那么$\{a * b_t\}$就是滤波后的输出结果, 即是$\{b_t\}$的滤波变换. 我们可以用如下所示表示滤波的一个流程图:

$$\{b_t\} \longrightarrow \boxed{\{a_t\}} \longrightarrow \{a * b_t\},$$

这说明把$\{b_t\}$经过$\{a_t\}$滤波可以得到$\{a * b_t\}$. 因为$\{a_t\} \leftrightarrow A(\cdot)$, 可以用$A(\cdot)$等价代替来用于滤波. 因此, 方便的是可以把一个滤波器作为抽象的结构, 它的两个表示就是$\{a_t\}$和$A(\cdot)$, 所以我们同样可以称"滤波器$\{a_t\}$"或"滤波器$A(\cdot)$". "时域"的表示式$\{a_t\}$称为滤波器的*脉冲响应序列*, 而"频域"的表示式$A(\cdot)$称为*传递函数*(或*频率响应函数*). 如下的流程图等价于上述流程图:

$$\{b_t\} \longrightarrow \boxed{A(\cdot)} \longrightarrow \{a * b_t\}.$$

因为一般来说滤波器传递函数$A(\cdot)$是复值的, 所以可以考虑它的极坐标表示, 即

$$A(f) = |A(f)| e^{i\theta(f)}.$$

这个由$|A(f)|$和$\theta(f)$定义的函数分别叫做滤波器的*增益函数*和*滤波相位函数*(注意$\theta(f)$仅当$|A(f)| > 0$时才有意义). 相似地, $|A(f)|^2$用于定义*平方增益函数*.

作为我们举出的第一个具体的滤波器的例子, 考虑脉冲响应序列

$$a_t = \begin{cases} 1/2, & t = 0; \\ 1/4, & t = \pm 1; \\ 0, & \text{其他} \end{cases} \tag{25d}$$

给出的滤波器.

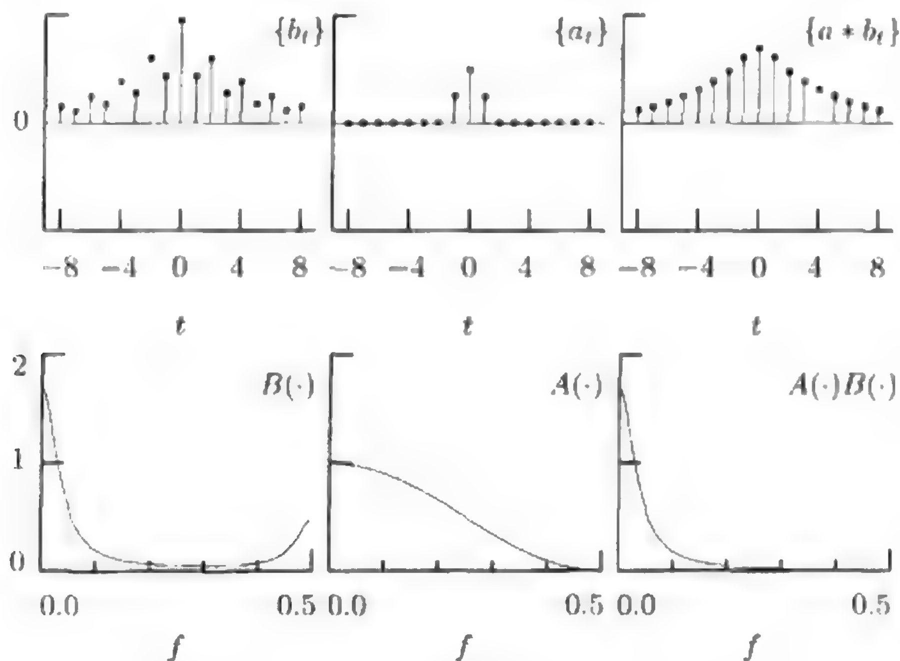


图 26 用低通滤波器进行滤波的例子

练习[2.3]指出它的传递函数 $A(\cdot)$是用 $\cos^2(\pi f)$给出. 我们需要一个无限序列来进行滤波，因此，考虑

$$b_t \equiv \sum_{l=1}^{L} C_l \phi_l^{|t|}, \quad t = \cdots, -1, 0, 1, \cdots, \tag{26a}$$

其中 C_l 与 ϕ_l 都是实值，且 $|\phi_l| < 1$. 练习[2.4]也表明$\{b_t\}$的离散傅里叶变换由下式给出：

$$B(f) \equiv \sum_{l=1}^{L} C_l \frac{1-\phi_l^2}{1-2\phi_l\cos(2\pi f)+\phi_l^2}. \tag{26b}$$

因此用$\{a_t\}$来滤波$\{b_t\}$的结果的离散傅里叶变换(DFT)由 $A(f)B(f)=\cos^2(\pi f)B(f)$ 给出.

图 26 中所示的例子表明在 $L=2$，$C_1=3/16$，$\phi_1=4/5$，$C_2=1/20$ 以及 $\phi_2=-4/5$ 的情况下(序列的图在 $t=\pm 8$ 时被截断). 离散傅里叶变换 $B(\cdot)$在 $f=0$ 处到达顶点并在 $f=1/2$ 处到达高频顶点. $\{b_t\}$的低频顶点在 $t=0$ 时达到最大值 19/80，并且当 $t\to\pm\infty$时持续衰退. 高频顶点是由相邻值间的“波浪般”变化形成的. $\{a_t\}$的传递函数指出这种滤波叫做“低通滤波”：它在削弱高频部分的同时保持低频部分不变. 滤波系列的离散傅里叶变换 $A(\cdot)B(\cdot)$在高频时非常小，反映在$\{a*b_t\}$中就是，和$\{b_t\}$相比，缺少“起伏”.

对照来看，让我们考虑脉冲响应序列为

$$a_t = \begin{cases} 1/2, & t=0; \\ -1/4, & t=\pm 1; \\ 0, & \text{其他} \end{cases} \tag{26c}$$

的滤波器.

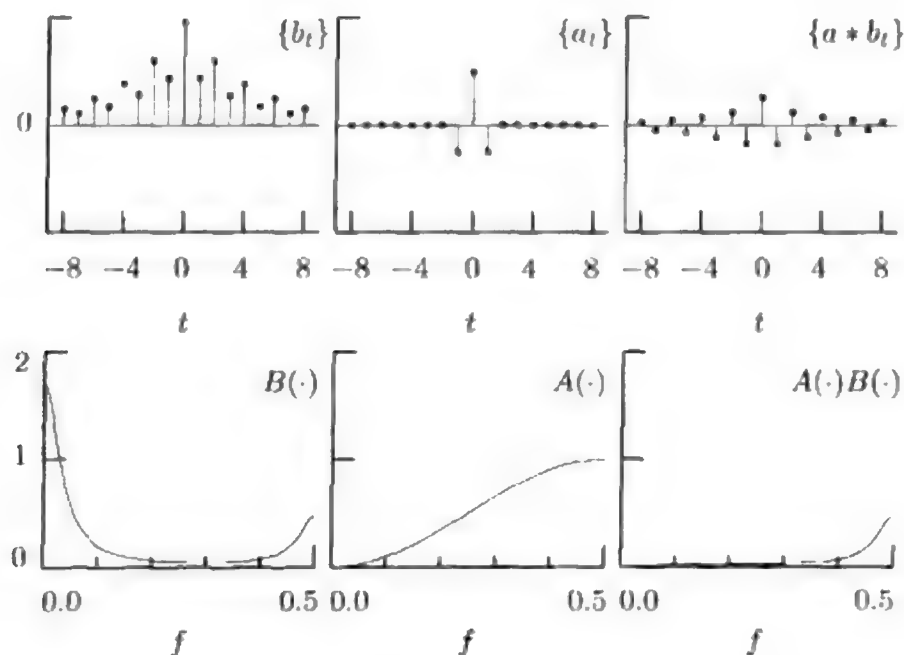


图 27 使用高通滤波器进行滤波的例子

练习[2.5]指出，它的传递函数由 $A(f)=\sin^2(\pi f)$ 给出. 这个传递函数如图 27下面一行的中间图所示，并且表明上面的是"高通"滤波，因为它在削弱低频部分的同时保持高频部分不变. 滤波系列 $\{a*b_t\}$ 由 $\{b_t\}$ 中在(25d)中被消去的波浪起伏般的部分组成.

最后让我们来介绍一下最重要的概念：**滤波器的串联**. 这种现象在我们学习小波的过程中会多次遇到. 滤波器的串联(级联)的定义就是：它是一组 M 个滤波的组合，即第一次滤波的输出就是第二次滤波的输入并持续进行.

练习[27] 假定 $\{a_{m,t}: t=\cdots, -1, 0, 1, \cdots\}$，$m=1, \cdots, M$，是和传递函数 $A_m(\cdot)$ 相对应的 M 个滤波器. 假定这些滤波器形成串联，它的第 m 次滤波器是 $\{a_{m,t}\}$. 将 $\{b_t\}$ 输入串联滤波器(即输入到 $\{a_{1,t}\}$)，且让输出为 $\{c_t\}$(即由 $\{a_{M,t}\}$ 输出). 我们可以得到流程图

$$\{b_t\} \longrightarrow \boxed{A_1(\cdot)} \longrightarrow \boxed{A_2(\cdot)} \longrightarrow \cdots \longrightarrow \boxed{A_M(\cdot)} \longrightarrow \{c_t\}.$$

说明我们可以得到

$$c_t = \sum_{u=-\infty}^{\infty} a_u b_{t-u}, \quad t = \cdots, -1, 0, 1, \cdots,$$

其中 $\{a_t\}$ 是一个滤波器，它的脉冲响应序列由单个滤波器连续卷积而成，其传递函数为

$$A(f) \equiv \prod_{m=1}^{M} A_m(f)$$

(回想我们的假设是一个脉冲响应序列，且由一个有限序列的平方和卷积形成).

◁

滤波器$\{a_t\}$和串联滤波器是等价的滤波器，以上的流程图可以重新写成

$$\{b_t\} \longrightarrow \boxed{A(\cdot)} \longrightarrow \{c_t\},$$

其中$c_t = a * b_t$.

练习[28a] 假设第m个滤波器为$a_{m,K_m} \neq 0$且$a_{m,K_m+L_m-1} \neq 0$，则对于所有的$t \leq K_m - 1$及$t \geq K_m + L_m$，有$a_{m,t} = 0$(假定L_m是一个正整数). 因为这种滤波器是由除零以外的L_m个相邻元素组成，我们就说它具有L_m的宽度. 它的传递函数是

$$A_m(f) = \sum_{t=K_m}^{K_m+L_m-1} a_{m,t} e^{-i2\pi ft}.$$

等价于$A(\cdot)$的滤波器宽度是

$$L \equiv \sum_{m=1}^{M} L_m - M + 1. \tag{28a}$$

◁

2.4 有限序列的傅里叶变换

现在，假设$\{a_t\} = \{a_t : t = 0, \cdots, N-1\}$是一个有$N$个实值或复值变量的序列. 它的离散傅里叶变换(DFT)是由N个变量组成的序列$\{A_k\}$

$$A_k \equiv \sum_{t=0}^{N-1} a_t e^{-i2\pi tk/N}, \quad k = 0, \cdots, N-1, \tag{28b}$$

注意到A_k是连带频率$f_k \equiv k/N$. 上式右边可以用于对所有整数的k定义A_k，它产生一个具有下述性质的无穷序列.

练习[28b] 证明，如果对于所有的k，A_k是由式(28b)的右边定义的序列，那么得到的无穷序列是以N项为周期的. 即对任何非零整数n和任何整数k，$0 \leq k \leq N-1$，我们有$A_{k+nN} = A_k$(注意这个和练习[22a]的类似). ◁

有些作者称$\{A_k\}$为“有限傅里叶变换”并且只对无穷序列使用“离散傅里叶变换”这个名字. 我们更喜欢用“离散傅里叶变换”来应用于有限与无穷序列. 使用“有限傅里叶变换”导出不适宜的缩写词“FFT”，现在普遍用于“快速傅里叶变换”算法，很快地计算$\{A_k\}$的方法的集合(见Oppenheim and Schafer，1989).

我们能够用如下方式从它的离散傅里叶变换$\{A_k\}$来重构$\{a_t\}$.

练习[29a] 证明

$$\frac{1}{N}\sum_{k=0}^{N-1}A_k e^{i2\pi tk/N}=a_t,\quad t=0,\cdots,N-1.$$

提示：如果 z 是任意一个不等于 1 的复数，那么

$$\sum_{u=0}^{N-1}z^u=\frac{1-z^N}{1-z},$$

并且之后注意到，当 $z=\exp(i2\pi/N)$时，我们可以写 $\exp(i2\pi u/N)=z^u$. ◁

因为我们可以通过$\{a_t\}$的离散傅里叶变换来重构$\{a_t\}$，那么这两个序列$\{a_t\}$和$\{A_k\}$能够考虑作为同一个数学实体的等价表示. $\{a_t\}$和$\{A_k\}$之间的傅里叶关系如

$$\{a_t\}\leftrightarrow\{A_k\},$$

也就是

$$A_k=\sum_{t=0}^{N-1}a_t e^{-i2\pi tk/N}$$

和

$$a_t=\frac{1}{N}\sum_{k=0}^{N-1}A_k e^{i2\pi tk/N},$$

其中$\{a_t\}$现在可以称作$\{A_k\}$的离散傅里叶变换的逆. 对 $t\leqslant-1$ 和 $t\geqslant N$，上面右式的和可以用来定义 a_t，得到一个以 N 为周期的无穷序列(这个结果是在练习[28b]基础上的一个变型).

对于有限序列的练习[23a]的类似结果如下.

练习[29b] 假定$\{a_t\}\leftrightarrow\{A_k\}$并且$\{b_t\}\leftrightarrow\{B_k\}$. 证明

$$\sum_{t=0}^{N-1}a_t b_t^*=\frac{1}{N}\sum_{k=0}^{N-1}A_k B_k^*.$$ ◁

令 $b_t=a_t$，我们得到对于有限序列的帕塞瓦尔定理，即

$$\sum_{t=0}^{N-1}|a_t|^2=\frac{1}{N}\sum_{k=0}^{N-1}|A_k|^2.$$

2.5 有限序列的循环卷积/滤波

我们再次假定$\{a_t\}$和$\{b_t\}$是长度为 N 的有限序列，使$\{a_t\}\leftrightarrow\{A_k\}$，$\{b_t\}\leftrightarrow\{B_k\}$. 让我们定义一个由$\{a_t\}$和$\{b_t\}$形成的卷积，它是一个 N 个元素的序列，定义为

$$a*b_t\equiv\sum_{u=0}^{N-1}a_u b_{t-u},\quad t=0,\cdots,N-1,$$

其中当 $t\leqslant -1$ 时，b_t 用周期延拓；即 $b_{-1}=b_{N-1}$，$b_{-2}=b_{N-2}$；以此类推．为了强调我们在使用周期延拓，可以把以上内容等价转换为

$$a * b_t \equiv \sum_{u=0}^{N-1} a_u b_{t-u \bmod N}, \quad t=0,\cdots,N-1,$$

其中"$t-u \bmod N$"表示"$t-u$ 以 N 为模"，定义如下：假设 j 是一个包含在 $0\leqslant j\leqslant N-1$ 内的整数，那么 $j \bmod N \equiv j$；假设 j 是另一个任意整数，那么，$j \bmod N \equiv j+nN$，其中 nN 是 N 的唯一整数倍，使 $0\leqslant j+nN\leqslant N-1$（一定要注意，当表示成 $b_{10+s+t-[(u+5v)/2] \bmod N}$ 这种形式时，求模的运算是我们计算 $10+s+t-[(u+5v)/2]$ 之后的结果）．然而，不管我们选择如何表达，上面的卷积称为*循环卷积*．图 31a 说明了对于这个术语的根本理由．

和无限序列的情况类似，循环卷积的离散傅里叶变换和 $\{a_t\}$ 与 $\{b_t\}$ 的离散傅里叶变换有非常简单的关系．

练习[30] 验证 $\{a * b_t\}$ 的离散傅里叶变换是用下式给出：

$$\sum_{t=0}^{N-1} a * b_t \mathrm{e}^{-\mathrm{i}2\pi tk/N} = A_k B_k, \quad k=0,\cdots,N-1;$$

即 $\{a * b_t\} \leftrightarrow \{A_k B_k\}$． ◁

有两个重要的相关的概念是复互相关，定义为

$$a^* \star b_t \equiv \sum_{u=0}^{N-1} a_u^* b_{u+t \bmod N}, \quad t=0,\cdots,N-1,$$

以及自相关，定义为

$$a^* \star a_t \equiv \sum_{u=0}^{N-1} a_u^* a_{u+t \bmod N}, \quad t=0,\cdots,N-1.$$

图 31b 给出了一个复互相关的图例．练习[2.9]指出 $\{a^* \star b_t\}$ 的离散傅里叶变换用

$$\sum_{t=0}^{N-1} a^* \star b_t \mathrm{e}^{-\mathrm{i}2\pi tk/N} = A_k^* B_k, \quad k=0,\cdots,N-1 \tag{30}$$

给出，即 $\{a^* \star b_t\} \leftrightarrow \{A_k^* B_k\}$．自相关序列的离散傅里叶变换直接为

$$\sum_{t=0}^{N-1} a^* \star a_t \mathrm{e}^{-\mathrm{i}2\pi tk/N} = A_k^* A_k = |A_k|^2, \quad k=0,\cdots,N-1,$$

即 $\{a^* \star a_t\} \leftrightarrow \{|A_k|^2\}$．

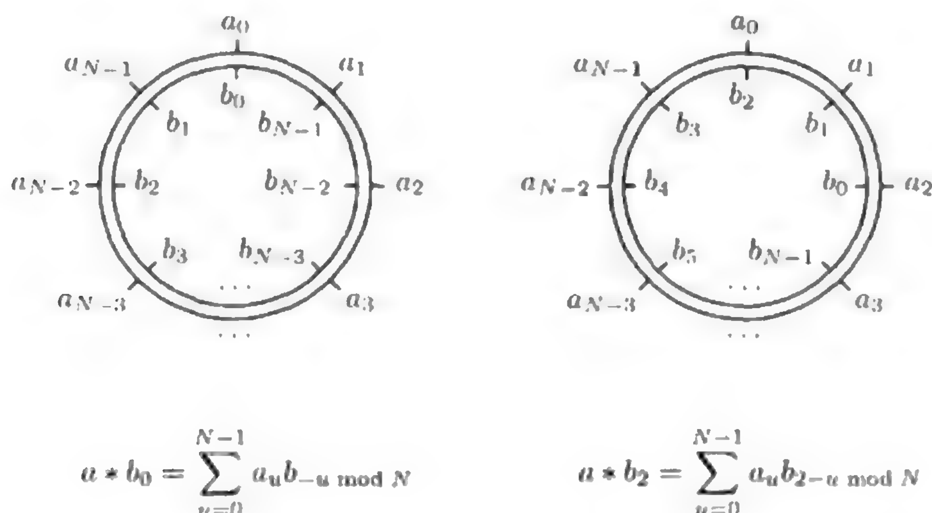


图 31a　有限序列$\{a_t\}$和$\{b_t\}$循环卷积的图例. 左图中两个环. 外部的环等分成 N 个弧, 弧的边界像钟表一样从上到下标记上 a_0, …, a_{N-1}. 内部的环也是如此等分, 不过现在逆时针标记上 b_0, …, b_{N-1}. $\{a_t\}$和$\{b_t\}$循环卷积的第零个元素 $a * b_0$ 由 a_t 和 b_t 中对应元素相乘然后相加得到. 总的来说, 第 t 个元素 $a * b_t$ 可以用相似的方法得到, 只不过内环是旋转钟到 t 份——举例来说, 右图说明如何用同心环队列得到循环卷积的第二个元素$a * b_2$

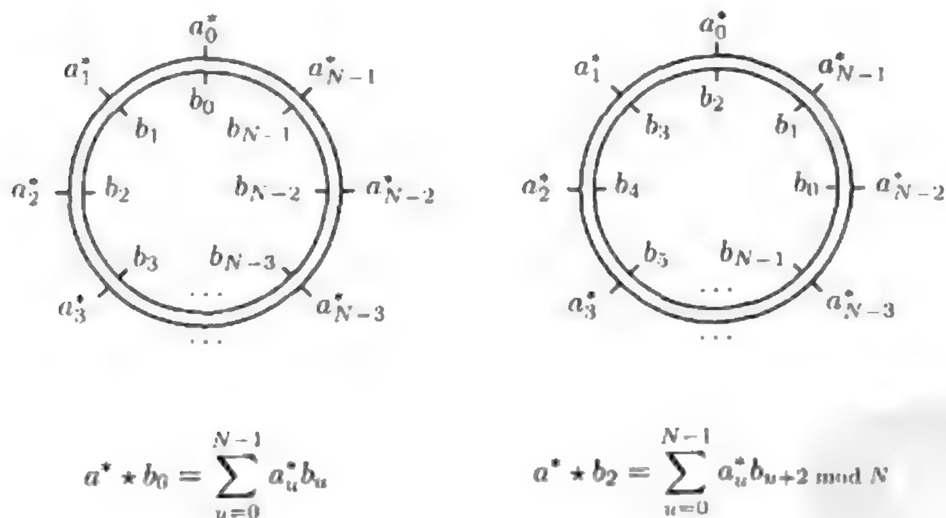


图 31b　$\{a_t\}$和$\{b_t\}$的循环复互相关的图例. 布局和图 31a 类似

练习[31]　证明, 适当地定义$\{\ddot{a}_t\}$, 使$\{a_t\}$和$\{b_t\}$复互相关得到的序列和$\{\ddot{a}_t\}$与$\{b_t\}$卷积得到的序列是相同的. 并证明如果$\{a_t\} \leftrightarrow \{A_k\}$, 那么$\{\ddot{a}_t\} \leftrightarrow \{A_k^*\}$; 即具有离散傅里叶变换$\{A_k\}$的序列和$\{b_t\}$的复互相关与具有离散傅里叶变换$\{A_k^*\}$的序列和$\{b_t\}$卷积是相同的. 提示: 对比图 31a 和 31b. ◁

在工程文献中循环滤波的概念和循环卷积的概念是相同的. 因此, 现在表示

$\{a_t\}$为一个循环滤波器并把$\{b_t\}$作为待滤波的有限序列，由滤波输出的结果是$\{a*b_t\}$，如前可以用如下流程图表示以上操作：

$$\{b_t\} \longrightarrow \boxed{\{a_t\}} \longrightarrow \{a*b_t\}. \tag{32a}$$

因为$\{a_t\}\leftrightarrow\{A_k\}$，对于循环滤波器的一个等价表示就是$\{A_k\}$，因此我们也可以如下表示流程图：

$$\{b_t\} \longrightarrow \boxed{\{A_k\}} \longrightarrow \{a*b_t\}.$$

而如文中所述，(32a)可以表示一个有限序列的循环滤波或者一个无限序列的滤波，以上仅涉及循环滤波：循环滤波器的离散傅里叶变换一定是一个和$\{A_k\}$类似的序列，而定义在所有整数上的滤波器的离散傅里叶变换一定是一个周期函数，即定义在实轴上的$A(\cdot)$.

循环滤波还能导出下一部分将要讨论的流程图的第三种表示方法.

2.6 周期滤波器

在小波学习的过程中，循环滤波器通常由一个包含无穷项的脉冲响应序列$\{a_t: t=\cdots, -1, 0, 1, \cdots\}$直接构建而成(我们将在第4章讨论离散Daubechies小波类，事实上除了有限项外，a_t其余全都为零)．给出有限序列$\{b_t: t=0, 1, \cdots, N-1\}$，让我们来构建有限序列

$$c_t \equiv \sum_{v=-\infty}^{\infty} a_v b_{t-v \bmod N}, \quad t = 0,1,\cdots,N-1. \tag{32b}$$

以上类似于一个循环滤波运算，由N个项组成的序列转换成一个长度为N的新序列，但是这种滤波器形式上有无穷多项．事实上，此处我们可以把以上运算重新表达为一个长度为N的循环滤波器，称为和$\{a_t\}$相对应的周期滤波器．为了完成这个运算，首先把无限多的和分成由N个邻接的数组成的有限和的组：

$$c_t = \sum_{n=-\infty}^{\infty} \sum_{u=nN}^{(n+1)N-1} a_u b_{t-u \bmod N} = \sum_{n=-\infty}^{\infty} \sum_{u=0}^{N-1} a_{u+nN} b_{t-(u+nN) \bmod N}.$$

因为对于所有的n，有$b_{t-(u+nN)\bmod N}=b_{t-u\bmod N}$，我们可以写为

$$c_t = \sum_{n=-\infty}^{\infty} \sum_{u=0}^{N-1} a_{u+nN} b_{t-u \bmod N} = \sum_{u=0}^{N-1} \Big(\sum_{n=-\infty}^{\infty} a_{u+nN} \Big) b_{t-u \bmod N}$$

$$\equiv \sum_{u=0}^{N-1} a_u^{\circ} b_{t-u \bmod N},$$

其中

$$a_t^\circ \equiv \sum_{n=-\infty}^{\infty} a_{t+nN}, \quad t=0,1,\cdots,N-1.$$

以上定义了一个相应于$\{a_t\}$的周期滤波器(也可以说$\{a_t^\circ\}$是$\{a_t\}$长度为 N 的周期化). 从本质上来说，滤波器$\{a_t\}$可以分割为长度为 N 的块

$$\cdots,\underbrace{a_{-N},a_{-N+1},\cdots,a_{-1}}_{\text{块}n=-1},\underbrace{a_0,a_1,\cdots,a_{N-1}}_{\text{块}n=0},\underbrace{a_N,a_{N+1},\cdots,a_{2N-1}}_{\text{块}n=1},\cdots$$

把这些项放平相加就形成周期滤波器波$\{a_t^\circ\}$:

$$\begin{array}{lcccc}
\vdots & \vdots & \vdots & \vdots & \vdots \\
 & + & + & \cdots & + \\
\text{块 } n=-1: & a_{-N} & a_{-N+1} & \cdots & a_{-1} \\
 & + & + & \cdots & + \\
\text{块 } n=0: & a_0 & a_1 & \cdots & a_{N-1} \\
 & + & + & \cdots & + \\
\text{块 } n=1: & a_N & a_{N+1} & \cdots & a_{2N-1} \\
 & + & + & \cdots & + \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
 & \Downarrow & \Downarrow & \cdots & \Downarrow \\
\text{周期滤波器:} & a_0^\circ & a_1^\circ & \cdots & a_{N-1}^\circ
\end{array}$$

现在令$\{A_k^\circ: k=0, \cdots, N-1\}$是$\{a_t^\circ\}$的离散傅里叶变换，并且让 $A(\cdot)$是对应于$\{a_t\}$的传递函数(即离散傅里叶变换). 以下练习表明$\{A_k^\circ\}$和 $A(\cdot)$在一定方式上是相关的.

练习[33] 证明

$$A_k^\circ = A\left(\frac{k}{N}\right). \qquad \triangleleft$$

因此，对于$\{a_t^\circ\}$的离散傅里叶变换$\{A_k^\circ\}$可以容易地从对应于初始滤波器的传递函数 $A(\cdot)$中得到；即$\{A_k^\circ\}$是由 $A(\cdot)$以 $1/N$ 为间隔进行取样得到的.

借助于流程图，我们可以表示式(32b)的(隐含的)循环滤波器为

$$\{b_t\} \longrightarrow \boxed{\left\{A\left(\frac{k}{N}\right)\right\}} \longrightarrow \{c_t\}.$$

为进一步参考，我们在这个流程图中做两个小的变化(参见以下的注释和扩展[1]). 首先，如果我们把序列$\{b_t\}$和$\{c_t\}$的元素表示成 N 维向量 $\mathbf{B}$ 和 $\mathbf{C}$，就可以得到

$$\mathbf{B} \longrightarrow \boxed{\left\{A\left(\frac{k}{N}\right)\right\}} \longrightarrow \mathbf{C}.$$

其次，有时候可以很方便地把$\left\{A\left(\frac{k}{N}\right)\right\}$(一个序列)缩写为 $A\left(\frac{k}{N}\right)$(序列中的一个元素):

$$\mathbf{B} \longrightarrow \boxed{A\left(\frac{k}{N}\right)} \longrightarrow \mathbf{C}. \qquad (33)$$

这种形式的流程图就是我们在后续几章中经常用到的形式.

作为产生一个周期滤波器的例子，考虑下述无穷序列：

$$a_t=\begin{cases}1/2, & t=0;\\ 1/4, & t=\pm 1;\\ 0, & \text{其他}.\end{cases}$$

现在这恰好是一个对应于式(25d)滤波器的脉冲响应序列的新记号，它的传递函数由练习[2.3]直接给出. 当 $N\geqslant 4$ 时，周期滤波器如下给出：

$$a_0^\circ=\sum_{n=-\infty}^{\infty}a_{nN}=a_0=1/2$$

$$a_1^\circ=\sum_{n=-\infty}^{\infty}a_{1+nN}=a_1=1/4$$

$$a_t^\circ=\sum_{n=-\infty}^{\infty}a_{t+nN}=0,\quad t=2,\cdots,N-2$$

$$a_{N-1}^\circ=\sum_{n=-\infty}^{\infty}a_{N-1+nN}=a_{-1}=1/4.$$

现在$\{a_t^\circ\}$的离散傅里叶变换按要求(见对于周期化的第二个例子的练习[2.11])如下：

$$A_k^\circ\equiv\sum_{t=0}^{N-1}a_t^\circ e^{-i2\pi tk/N}=\sum_{t=-1}^{1}a_{t\bmod N}^\circ e^{-i2\pi tk/N}=\sum_{t=-1}^{1}a_t e^{-i2\pi tk/N}$$

$$=\sum_{t=-\infty}^{\infty}a_t e^{-i2\pi tk/N}=A\left(\frac{k}{N}\right).$$

最后我们注意到一个关于滤波器$\{a_t\}$和它的周期化$\{a_t^\circ\}$的一个重要的结论.

练习[34] 假定滤波器$\{a_t\}$是已规范的，即

$$\sum_{t=-\infty}^{\infty}|a_t|^2=1.$$

如果$\{a_t^\circ\}$是$\{a_t\}$按长度 N 周期化的，证明一般来说

$$\sum_{t=0}^{N-1}|a_t^\circ|^2\neq 1;$$

即，一个滤波器的平方和与它的周期化的平方和不必是相同的.（提示：或者构造一个简单的反例或者对于$\{a_t\}$和$\{a_t^\circ\}$的离散傅里叶变换的关系使用帕塞瓦尔定理.） ◁

2.6 节的评论与扩展

[1]对于式(32b)的滤波运算来说，一个可能的可选择方法就是用式(33)的流程图来代替

$$\boldsymbol{B} \longrightarrow \boxed{A(\cdot)} \longrightarrow \boldsymbol{C}.$$

读者应仔细注意上式如何解释：$\boldsymbol{B}$ 的 N 个元素是由周期化滤波器循环滤波得到的，这个周期化滤波的频域(傅里叶)表示是 $\left\{A\left(\frac{k}{N}\right): k=0, \cdots, N-1\right\}$，滤波器的输出结果由 $\boldsymbol{C}$ 中的 N 个元素表示. 当以上流程图在小波文献中非常普遍时，$A(\cdot)$的取样就是隐含的，因此我们更愿意使用式(33)，因为它清楚地表示了取样运算.

[2]傅里叶理论中一个众所周知的现象是，时域内离散取样可以用频域范围内等分频带后得到的节点的和来表示(参见，例如，3.8 节，Percival and Walden，1993). 注意到周期化的滤波器这里可以被认为是"双重的"混淆现象，在这种现象中，时间和频率的角色互换了：等分的时域变量的和可以用频域的离散取样来表示.

2.7 傅里叶理论小结

稍后提及，我们从无穷序列的傅里叶变换中总结出最重要的结果，在实轴上的有限序列和实轴上定义的函数.

• 无限实值或复值序列$\{a_t\}$，并且 $\sum_{t=-\infty}^{\infty} |a_t|^2 < \infty$

(a)傅里叶表示：

$$a_t = \int_{-1/2}^{1/2} A(f) \mathrm{e}^{\mathrm{i}2\pi ft} \mathrm{d}f, \quad t = \cdots, -1, 0, 1, \cdots \tag{35a}$$

(重构公式)，其中

$$A(f) \equiv \sum_{t=-\infty}^{\infty} a_t \mathrm{e}^{-\mathrm{i}2\pi ft}, \quad |f| \leqslant 1/2 \tag{35b}$$

(分析公式). 上式右边可以用于对所有的 f 定义 $A(\cdot)$，得到一个有单位周期的周期函数. $\{a_t\}$和 $A(\cdot)$的傅里叶关系为

$$\{a_t\} \leftrightarrow A(\cdot).$$

(b)帕塞瓦尔定理："一个序列"的形式是

$$\sum_{t=-\infty}^{\infty} |a_t|^2 = \int_{-1/2}^{1/2} |A(f)|^2 \mathrm{d}f; \tag{35c}$$

而$\{b_t\} \leftrightarrow B(\cdot)$，"两个序列"的形式是

$$\sum_{t=-\infty}^{\infty} a_t^* b_t = \int_{-1/2}^{1/2} A^*(f) B(f) \mathrm{d}f. \tag{35d}$$

(c)卷积和相关理论：卷积

$$a * b_t \equiv \sum_{u=-\infty}^{\infty} a_u b_{t-u}, \quad t = \cdots, -1, 0, 1, \cdots \tag{36a}$$

的傅里叶变换是

$$\sum_{t=-\infty}^{\infty} a * b_t \mathrm{e}^{-\mathrm{i}2\pi ft} = A(f)B(f), \tag{36b}$$

因此$\{a * b_t\} \leftrightarrow A(\cdot)B(\cdot)$. 复互相关

$$a^* \star b_t \equiv \sum_{u=-\infty}^{\infty} a_u^* b_{u+t}, \quad t = \cdots, -1, 0, 1, \cdots \tag{36c}$$

的傅里叶变换是

$$\sum_{t=-\infty}^{\infty} a^* \star b_t \mathrm{e}^{-\mathrm{i}2\pi ft} = A^*(f)B(f), \tag{36d}$$

因此，$\{a^* \star b_t\} \leftrightarrow A^*(\cdot)B(\cdot)$. 令 $b_t = a_t$，得到自相关，对于自相关

$$\{a^* \star a_t\} \leftrightarrow |A(\cdot)|^2. \tag{36e}$$

• *有限实值或复值序列*$\{a_t : t=0, \cdots, N-1\}$

(a)傅里叶表示：

$$a_t = \frac{1}{N}\sum_{k=0}^{N-1} A_k \mathrm{e}^{\mathrm{i}2\pi tk/N}, \quad t = 0, \cdots, N-1 \tag{36f}$$

(重构公式)，其中

$$A_k \equiv \sum_{t=0}^{N-1} a_t \mathrm{e}^{-\mathrm{i}2\pi tk/N}, \quad k = 0, \cdots, N-1 \tag{36g}$$

(分析公式). 注意到 A_k 是和频率 $f_k \equiv k/N$ 联系起来的. 上面两个公式的右边可以用于对所有的整数 t 和 k 定义 a_t 和 A_k，得到周期为 N 的周期无限序列. $\{A_k\}$ 和$\{a_t\}$的傅里叶关系为：

$$\{a_t\} \leftrightarrow \{A_k\}.$$

(b)帕塞瓦尔定理："一个序列"的形式是

$$\sum_{t=0}^{N-1} |a_t|^2 = \frac{1}{N}\sum_{k=0}^{N-1} |A_k|^2; \tag{36h}$$

而$\{b_t\} \leftrightarrow \{B_k\}$，"两个序列"的形式是

$$\sum_{t=0}^{N-1} a_t^* b_t = \frac{1}{N}\sum_{k=0}^{N-1} A_k^* B_k. \tag{36i}$$

(c)卷积和相关定理：卷积

$$a * b_t \equiv \sum_{u=0}^{N-1} a_u b_{t-u \bmod N}, \quad t = 0, \cdots, N-1 \tag{37a}$$

的傅里叶变换是

$$\sum_{t=0}^{N-1} a * b_t \mathrm{e}^{-\mathrm{i}2\pi tk/N} = A_k B_k, \quad k = 0, \cdots, N-1, \tag{37b}$$

因此$\{a * b_t\} \leftrightarrow \{A_k B_k\}$. 这种形式的卷积称为循环卷积或者周期卷积. 复相关

$$a^* \star b_t \equiv \sum_{u=0}^{N-1} a_u^* b_{u+t \bmod N}, \quad t = 0, \cdots, N-1 \tag{37c}$$

的傅里叶变换是

$$\sum_{t=0}^{N-1} a^* \star b_t \mathrm{e}^{-\mathrm{i}2\pi tk/N} = A_k^* B_k, \quad k = 0, \cdots, N-1, \tag{37d}$$

因此 $\{a^* \star b_t\} \leftrightarrow \{A_k^* B_k\}$. 令 $b_t = a_t$ 产生自相关，对于它来说

$$\{a^* \star a_t\} \leftrightarrow \{|A_k|^2\}. \tag{37e}$$

如果$\{a_t : t = \cdots, -1, 0, 1, \cdots\} \leftrightarrow A(\cdot)$，那么

$$c_t \equiv \sum_{u=-\infty}^{\infty} a_u b_{t-u \bmod N}, \quad t = 0, \cdots, N-1$$

能够再表示为一个循环卷积：

$$c_t = \sum_{u=0}^{N-1} a_u^\circ b_{t-u \bmod N},$$

其中 $a_t^\circ \equiv \sum_{u=-\infty}^{\infty} a_{t+uN}$. 长度为 N 的序列$\{a_t^\circ\}$是由长度为 N 的$\{a_t\}$周期化形成的，我们有

$$\{a_t^\circ : t = 0, \cdots, N-1\} \leftrightarrow \left\{A\left(\frac{k}{N}\right) : k = 0, \cdots, N-1\right\}.$$

- 实值或复值函数 $a(\cdot)$，并且 $\int_{-\infty}^{\infty} |a(t)|^2 \mathrm{d}t < \infty$

(a)傅里叶表示：

$$a(t) \overset{\mathrm{ms}}{=} \int_{-\infty}^{\infty} A(f) \mathrm{e}^{\mathrm{i}2\pi ft} \mathrm{d}f, \quad -\infty < t < \infty \tag{37f}$$

(重构公式)，其中

$$A(f) \equiv \int_{-\infty}^{\infty} a(t) \mathrm{e}^{-\mathrm{i}2\pi ft} \mathrm{d}t, \quad -\infty < t < \infty \tag{37g}$$

(分析公式). 记号"$\stackrel{\text{ms}}{=}$"表示"在平方意义下相等", 也就是说, (37f)等式左边和等式右边不一定对所有给定的 t 都要相等, 我们有

$$\int_{-\infty}^{\infty}\left|a(t)-\int_{-\infty}^{\infty}A(f)\mathrm{e}^{\mathrm{i}2\pi ft}\mathrm{d}f\right|^2\mathrm{d}t=0.$$

$a(\cdot)$和 $A(\cdot)$之间的傅里叶关系记为

$$a(\cdot)\leftrightarrow A(\cdot).$$

(b)帕塞瓦尔定理: "一个函数"的形式是

$$\int_{-\infty}^{\infty}|a(t)|^2\mathrm{d}t=\int_{-\infty}^{\infty}|A(f)|^2\mathrm{d}f; \tag{38a}$$

而 $b(\cdot)\leftrightarrow B(\cdot)$, "两个函数"的形式是

$$\int_{-\infty}^{\infty}a^*(t)b(t)\mathrm{d}t=\int_{-\infty}^{\infty}A^*(f)B(f)\mathrm{d}f. \tag{38b}$$

(c)卷积和相关定理: 卷积

$$a*b(t)\equiv\int_{-\infty}^{\infty}a(u)b(t-u)\mathrm{d}u \tag{38c}$$

的傅里叶变换是

$$\int_{-\infty}^{\infty}a*b(t)\mathrm{e}^{-\mathrm{i}2\pi ft}\mathrm{d}t=A(f)B(f), \tag{38d}$$

因此 $a*b(\cdot)\leftrightarrow A(\cdot)B(\cdot)$. 复相关

$$a^*\star b(t)\equiv\int_{-\infty}^{\infty}a^*(u)b(u+t)\mathrm{d}u \tag{38e}$$

的傅里叶变换是

$$\int_{-\infty}^{\infty}a^*\star b(t)\mathrm{e}^{-\mathrm{i}2\pi ft}\mathrm{d}t=A^*(f)B(f), \tag{38f}$$

因此 $a^*\star b(\cdot)\leftrightarrow A^*(\cdot)B(\cdot)$. 令 $b(f)=a(f)$产生一个自相关, 即

$$a^*\star a(\cdot)\leftrightarrow|A(\cdot)|^2. \tag{38g}$$

2.8 练习

[2.1] 假设$\{a_t: t=\cdots, -1, 0, 1, \cdots\}$是一个实值序列，它的离散傅里叶变换是$A(\cdot)$. 对于某个$f_0$使$0<f_0<1/2$，构造一个复值序列$c_t=a_t e^{-i2\pi f_0 t}$（这个程序称为复解调）. $\{c_t\}$的离散傅里叶变换$C(\cdot)$如何和$\{a_t\}$的离散傅里叶变换$A(\cdot)$建立联系？假设ε代表一个变量使$0<f_0-\varepsilon<f_0+\varepsilon<1/2$，并且$A(\cdot)$中的正频率能够和一些物理现象联系起来. 你可以把$C(f)$的负频率$-\varepsilon\leqslant f<0$作为物理现象中的哪种解释？

[2.2] 证明无限序列$\{a_t\}$和$\{b_t\}$的复相关$\{a^* \star b_t\}$的离散傅里叶变换由式(25a)给出.

[2.3] 证明滤波器的传递函数$A(\cdot)$是以$\cos^2(\pi f)$的形式给出的，其中滤波器的脉冲响应序列如式(25d)所示.

[2.4] 证明：由式(26a)定义的序列$\{b_t\}$的离散傅里叶变换可以由式(26b)定义的函数$B(\cdot)$给出. 提示：当$z\equiv\phi_t e^{-i2\pi f}$时，$\phi_t^t e^{-i2\pi ft}$写为$z^t$.

[2.5] 证明具有式(26c)给出的频率脉冲响应的滤波器的传递函数$A(\cdot)$由$\sin^2(\pi f)$给出.

[2.6] 校验图26和27的内容.

[2.7] 证明由式(25d)定义的低通滤波器等价于由两个滤波器$\{a_{1,t}\}$和$\{a_{2,t}\}$的串联滤波器，和这两个的滤波宽度相同. 确定这些滤波器的传递函数$A_1(\cdot)$和$A_2(\cdot)$，且验证$A_1(f)A_2(f)=\cos^2(\pi f)$.

[2.8] 如果$\{a_t: t=0, \cdots, N-1\}\leftrightarrow A(\cdot)$，那么$a_{-1}$和$a_{N+1}$取什么值？如果$\{a_t: t=0, \cdots, N-1\}\leftrightarrow\{A_k\}$，那么$a_{-1}$和$a_{N+1}$取什么值？

[2.9] 证明有限序列$\{a_t\}$和$\{b_t\}$的循环复互相关$\{a^* \star b_t\}$的离散傅里叶变换如式(30)所示.

[2.10] 令$\{a_t: t=\cdots, -1, 0, 1, \cdots\}$是一个无穷序列，它的离散傅里叶变换是$A(\cdot)$. 求出无穷序列$\{a_{2n+1}: n=\cdots, -1, 0, 1, \cdots\}$（即$\{a_t\}$中的奇数项）借助于$A(\cdot)$表示的离散傅里叶变换.（练习[23b]给出了偶数项对应的傅里叶变换.）

[2.11] 令$\{b_t: t=0, \cdots, 3\}$是一个长度为4的有限序列，考虑一个由$a_t=\phi^{|t|}$，$t=\cdots, -1, 0, 1, \cdots$定义的滤波器，其中$\phi$是实值变量并满足$|\phi|<1$. 令

$$c_t\equiv\sum_{u=-\infty}^{\infty}a_u b_{t-u \bmod 4},\quad t=0,\cdots,3$$

表示$\{b_t\}$用$\{a_t\}$滤波后的结果. 由$\{a_t\}$设计长度为4的周期滤波器，即滤波器$\{a_t^\circ: t=0, \cdots, 3\}$使

$$c_t=\sum_{u=0}^{3}a_u^\circ b_{t-u \bmod 4}.$$

确定$\{a_t\}$的离散傅里叶变换$A(\cdot)$以及$\{a_t^\circ\}$的离散傅里叶变换$\{A_k^\circ\}$，当

$k=0,\cdots,3$时，验证 $A_k^\circ = A\left(\frac{k}{4}\right)$.

[2.12] 考虑以下定义在整数上的滤波器：

$$a_t = \begin{cases} 1/2, & t = 0,1,2 \text{ 或 } 3; \\ 0, & t < 0 \text{ 或 } t \geqslant 4. \end{cases}$$

相应于长度为 $N=1, 2, 3, 4$ 和 5 的周期滤波器 $\{a_t^\circ\}$ 是什么？

对于滤波器

$$a_t = \begin{cases} 1/2, & t = 0,1; \\ -1/2, & t = 2,3; \\ 0, & t < 0 \text{ 或 } t \geqslant 4 \end{cases}$$

回答同样的问题.

（第 4 章之后，我们能够将以上滤波器分别解释为 2 级水平的哈尔尺度和小波滤波器.）

第 3 章

时间序列的规范正交变换

3.0 引言

时间序列的小波分析可以用规范正交变换的形式来定义，因此这里我们简略地回顾一下这些变换背后的主要思想．首先在 3.1 节回顾一下规范正交变换的基本理论．3.2 节讨论重要的投影定理，而在 3.3 节考虑复值变换．在第 4 章引入离散小波变换(DWT)之前，我们先在 3.4 节讨论规范正交离散傅里叶变换(ODFT)，这是因为它在许多有趣的方面和离散小波变换是并行和形成对照的．3.5 节总结本章的要点，已经对规范正交变换比较熟悉的读者可以简单地阅读这一节来熟悉我们的符号和约定．

3.1 规范正交变换的基本理论

规范正交变换是非常有趣的变换，这是因为它们可以以某种方式重新表示时间序列，这种表示方式使得我们可以很容易地从它的变换重构时间序列．在不严格的意义下，变换中的“信息”和原来序列中的“信息”是等价的；换一种说法，序列和它的变换可以被认为是同一种数学实体的两种表达．规范正交变换以一种标准化的形式(例如傅里叶级数)来重新表示序列，以利于进一步的处理，以此将序列简化成概述其主要特征的一些数值(压缩)，并且分析序列来探索感兴趣的特殊模式(例如方差分析)．在文献中有很多规范正交变换的例子，包括离散傅里叶变换的规范正交形式(见 3.4 节)、离散余弦变换的不同形式(见练习[3.3])、沃尔什(Walsh)变换(Beauchamp，1984)，以及我们将在第 4 章中见到的离散小波变换．

令 X_0，X_1，…，X_{N-1}表示 N 个实值变量的时间序列(以后我们将把这样的序列记为$\{X_t: t=0, \cdots, N-1\}$，如果哑标 t 的给定值是清楚的，也可以记为$\{X_t\}$)．图 42 展示了两个这样的 $N=16$ 的序列．令 $\mathbf{X}$ 表示 N 维列向量，它的第 t 个元素是 X_t，$t=0, 1, \cdots, N-1$．如果 $\mathbf{Y}$ 是包含 Y_0，…，Y_{N-1}的另一个这样

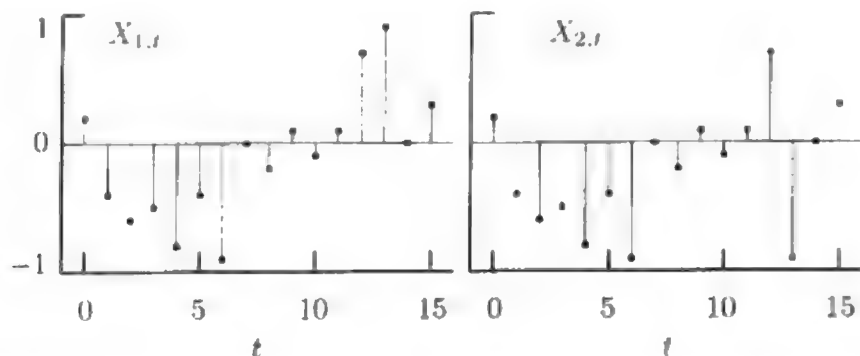


图 42 两个小的时间序列$\{X_{1,t}\}$和$\{X_{2,t}\}$，每个有 16 个值．左图用方块点表示$X_{1,t}$在$t=0,\cdots,15$处的值．右图表示$X_{2,t}$．两个时间序列仅仅在$t=13$处不同，$X_{2,13}=-X_{1,13}$．对于这个图示，$\{X_{1,t}\}$的 16 个值是 0.2，−0.4，−0.6，−0.5，−0.8，−0.4，−0.9，0.0，−0.2，0.1，−0.1，0.1，0.7，0.9，0.0 和 0.3

的向量，那么 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的内积由下式给出：

$$\langle \boldsymbol{X},\boldsymbol{Y}\rangle \equiv \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} = \sum_{t=0}^{N-1} X_t Y_t,$$

这里 $\boldsymbol{X}^{\mathrm{T}}$ 是 $\boldsymbol{X}$ 的转置．$\boldsymbol{X}$ 范数的平方由

$$\|\boldsymbol{X}\|^2 \equiv \langle \boldsymbol{X},\boldsymbol{X}\rangle = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \sum_{t=0}^{N-1} X_t^2$$

给出．我们用量 $\varepsilon_X \equiv \|\boldsymbol{X}\|^2$ 表示时间序列$\{X_t\}$的能量（ε 被物理学家称为能量，必须以 ε 的某种单位来度量 X_t）．

令$\mathcal{O}$表示一 $N\times N$ 的实值矩阵，满足规范正交性质$\mathcal{O}^{\mathrm{T}}\mathcal{O}=I_N$，这里 I_N 是 $N\times N$单位矩阵．对于 j，$k=0$，1，…，$N-1$，令$\mathcal{O}_{j\cdot}^{\mathrm{T}}$和$\mathcal{O}_{\cdot k}$分别表示$\mathcal{O}$的第 j 个行向量和第 k 个列向量．例如，

$$\text{如果}\mathcal{O} = \begin{bmatrix} 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \\ 1 & 0 & 0 \end{bmatrix},\text{那么}\mathcal{O}_{2\cdot} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\mathcal{O}_{\cdot 2} = \begin{bmatrix} \sin\theta \\ \cos\theta \\ 0 \end{bmatrix}.$$

利用这个记号，我们可以重写

$$\mathcal{O} = \begin{bmatrix} \mathcal{O}_{0\cdot}^{\mathrm{T}} \\ \mathcal{O}_{1\cdot}^{\mathrm{T}} \\ \vdots \\ \mathcal{O}_{N-1\cdot}^{\mathrm{T}} \end{bmatrix} = [\mathcal{O}_{\cdot 0},\mathcal{O}_{\cdot 1},\cdots,\mathcal{O}_{\cdot N-1}].$$

规范正交性质可以用内积的形式重新表示为

$$\langle \mathcal{O}_{\cdot k},\mathcal{O}_{\cdot k'}\rangle = \delta_{k,k'} \equiv \begin{cases} 1, & \text{如果 } k=k'; \\ 0, & \text{其他}, \end{cases}$$

这里$\delta_{j,k}$表示克罗内克函数. 规范正交性隐含了矩阵$\mathcal{O}$的逆$\mathcal{O}^{-1}$就是它的转置$\mathcal{O}^{\mathrm{T}}$. 这样$\mathcal{O}\mathcal{O}^{-1}=\mathcal{O}\mathcal{O}^{\mathrm{T}}=I_N$，由此可以推出$\langle \mathcal{O}_{j\cdot}, \mathcal{O}_{j'\cdot}\rangle=\delta_{j,j'}$. 所以$\mathcal{O}$的列向量和行向量都形成一个规范正交向量集.

我们可以将时间序列$\{X_t\}$关于规范正交正交矩阵$\mathcal{O}$分解，通过在$\boldsymbol{X}$左乘矩阵$\mathcal{O}$得到

$$\boldsymbol{O}\equiv\mathcal{O}\,\boldsymbol{X}=\begin{bmatrix}\mathcal{O}_{0\cdot}^{\mathrm{T}}\\ \mathcal{O}_{1\cdot}^{\mathrm{T}}\\ \vdots\\ \mathcal{O}_{N-1\cdot}^{\mathrm{T}}\end{bmatrix}\boldsymbol{X}=\begin{bmatrix}\mathcal{O}_{0\cdot}^{\mathrm{T}}\boldsymbol{X}\\ \mathcal{O}_{1\cdot}^{\mathrm{T}}\boldsymbol{X}\\ \vdots\\ \mathcal{O}_{N-1\cdot}^{\mathrm{T}}\boldsymbol{X}\end{bmatrix}=\begin{bmatrix}\langle\boldsymbol{X},\mathcal{O}_{0\cdot}\rangle\\ \langle\boldsymbol{X},\mathcal{O}_{1\cdot}\rangle\\ \vdots\\ \langle\boldsymbol{X},\mathcal{O}_{N-1\cdot}\rangle\end{bmatrix},\tag{43a}$$

这里我们用到了这样的事实，即$\mathcal{O}_{j\cdot}^{\mathrm{T}}\boldsymbol{X}=\langle\mathcal{O}_{j\cdot},\boldsymbol{X}\rangle=\langle\boldsymbol{X},\mathcal{O}_{j\cdot}\rangle$. N维列向量$\boldsymbol{O}$称为$\boldsymbol{X}$关于规范正交变换$\mathcal{O}$的变换系数. 第j个变换系数是O_j，以内积$\langle\boldsymbol{X},\mathcal{O}_{j\cdot}\rangle$给出. 如果我们在上面等式的两边再左乘$\mathcal{O}^{\mathrm{T}}$，并且注意到$\mathcal{O}^{\mathrm{T}}\mathcal{O}=I_N$，就可以得到重构等式

$$\boldsymbol{X}=\mathcal{O}^{\mathrm{T}}\boldsymbol{O}=[\mathcal{O}_{0\cdot},\mathcal{O}_{1\cdot},\cdots,\mathcal{O}_{N-1\cdot}]\begin{bmatrix}O_0\\ O_1\\ \vdots\\ O_{N-1}\end{bmatrix}=\sum_{j=0}^{N-1}O_j\mathcal{O}_{j\cdot},\tag{43b}$$

这告诉我们如何从它的变换系数$\boldsymbol{O}$重构$\boldsymbol{X}$（记号$O_j\mathcal{O}_{j\cdot}$是指用向量$\mathcal{O}_{j\cdot}$的每个元素乘标量O_j得到的向量）. 因为$O_j=\langle\boldsymbol{X},\mathcal{O}_{j\cdot}\rangle$，我们可以将$\boldsymbol{X}$的重构公式重新表示为

$$\boldsymbol{X}=\sum_{j=0}^{N-1}\langle\boldsymbol{X},\mathcal{O}_{j\cdot}\rangle\,\mathcal{O}_{j\cdot}.\tag{43c}$$

上式对任意的$\boldsymbol{X}$都成立，因此向量$\mathcal{O}_{j\cdot}$形成N维实值向量空间$\mathbf{R}^N$的一个基底；即任何N维实值列向量都可以表示成$\mathcal{O}_{0\cdot}$，$\mathcal{O}_{1\cdot}$，…，$\mathcal{O}_{N-1\cdot}$的一个唯一的线性组合.

练习[43] 证明对于$\mathcal{O}_{\cdot k}$存在类似的陈述，即

$$\boldsymbol{X}=\sum_{k=0}^{N-1}\langle\boldsymbol{X},\mathcal{O}_{\cdot k}\rangle\,\mathcal{O}_{\cdot k}.\tag{43d}$$

换句话说，如果$\mathcal{O}$是一个$N\times N$的规范正交矩阵，那么任意的一个N维向量$\boldsymbol{X}$可以写为$\mathcal{O}$的N个列向量的线性组合. ◁

关于规范正交变换的一个重要的事实是保持能量，即变换系数$\boldsymbol{O}$中的能量和原来序列$\boldsymbol{X}$的能量相等，这可以从以下的讨论中看出：

$$\varepsilon_{\boldsymbol{O}}\equiv\|\boldsymbol{O}\|^2=\boldsymbol{O}^{\mathrm{T}}\boldsymbol{O}=(\mathcal{O}\,\boldsymbol{X})^{\mathrm{T}}\mathcal{O}\boldsymbol{X}=\boldsymbol{X}^{\mathrm{T}}\mathcal{O}^{\mathrm{T}}\mathcal{O}\boldsymbol{X}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}=\|\boldsymbol{X}\|^2=\varepsilon_{\boldsymbol{X}}$$

（一个保持能量的变换有时也叫做等度变换）. 因为$\varepsilon_{\boldsymbol{X}}=\sum X_t^2$，和$\varepsilon_{\boldsymbol{O}}=\sum O_j^2$，我们可以讨论序列$\{X_t^2\}$描述了时间序列的能量是如何通过时间分解的，同时序列$\{O_j^2\}$描述了能量是如何通过具有不同指标的变换系数分解的.

练习[44a] 利用式(43b)，证明

$$\|\boldsymbol{X}\|^2 = \sum_{j=0}^{N-1} \|O_j\mathcal{O}_{j\cdot}\|^2;$$

即 $\boldsymbol{X}$ 的平方范数等于重构 $\boldsymbol{X}$ 的所有向量 $O_j\mathcal{O}_{j\cdot}$ 的平方范数之和. ◁

尽管有无数种可能的规范正交变换，但是具有特殊兴趣的规范正交变换(如离散小波变换和规范正交离散傅里叶变换)则是那些我们可以把某些物理意义附于其变换系数 O_j 或重构向量 $O_j\mathcal{O}_{j\cdot}$ 的规范正交变换.

3.2 投影定理

规范正交变换的一个重要性质可以用投影定理来概括，这个定理可以阐述如下. 假设我们希望用向量 $\mathcal{O}_{0\cdot}$，$\mathcal{O}_{1\cdot}$，…，$\mathcal{O}_{N'-1\cdot}$ 的一个线性组合来逼近 $\boldsymbol{X}$，这里 $N'<N$，对于系数 $\{\alpha_j\}$ 的某些集合，$\boldsymbol{X}$ 的近似 $\hat{\boldsymbol{X}}$ 就可以表示成如下的形式：

$$\hat{\boldsymbol{X}} = \sum_{j=0}^{N'-1} \alpha_j \mathcal{O}_{j\cdot}.$$

进而，假定我们是基于最小二乘准则来选择系数 α_j 的，即使得差向量 $\boldsymbol{e}\equiv\boldsymbol{X}-\hat{\boldsymbol{X}}$ 在 $\|\boldsymbol{e}\|^2$ 对于系数 $\{\alpha_j\}$ 的所有可能的选择应该最小化的意义下尽可能的小. 投影定理说明选取 $\alpha_j=O_j=\langle\boldsymbol{X},\mathcal{O}_{j\cdot}\rangle$ 就可以满足最小二乘准则.

练习[44b] 验证在建立投影定理中的关键一步，即

$$\|\boldsymbol{e}\|^2 = \sum_{j=0}^{N'-1}(O_j-\alpha_j)^2 + \sum_{j=N'}^{N-1}O_j^2.$$ ◁

注意到第一个平方和项可以通过令 $\alpha_j=O_j$ 来最小化(由于第二个平方和项所涉及的 $\mathcal{O}_{j\cdot}$ 并没有用来形成逼近 $\hat{\boldsymbol{X}}$，所以我们无法控制)，通过上式可以得到投影定理.

通过最小二乘选择的 α_j 即 $\alpha_j=O_j$，$j=0, 1, \cdots, N'-1$，我们有

$$\|\boldsymbol{e}\|^2 = \sum_{j=N'}^{N-1}O_j^2,$$

这是一个有用的刻画 $\hat{\boldsymbol{X}}$ 逼近于 $\boldsymbol{X}$ 的程度的度量. 注意到增加 N' 可以导致 $\|\boldsymbol{e}\|^2$ 显著地减小，即用更多的 $\mathcal{O}_{j\cdot}$ 来逼近 $\boldsymbol{X}$ 不会引起 $\|\boldsymbol{e}\|^2$ 的增加.

最后让我们回顾一下下述线性代数中的概念，这些概念将会在第4章中用到.

由于所有的 N 个向量$\mathcal{O}_{j\cdot}$构成了有限维空间 $\mathbf{R}^N$ 的一个基，故向量$\mathcal{O}_{j\cdot}$ $(0\leqslant j\leqslant N'-1$，$N'<N)$的集合组成了 $\mathbf{R}^N$ 子空间的一个基. 我们可以认为这个子空间是由向量$\mathcal{O}_{0\cdot}$，…，$\mathcal{O}_{N'-1\cdot}$张成的. 例如

$$\text{如果}\mathcal{O}=\begin{bmatrix}0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ 1 & 0 & 0\end{bmatrix},\quad \text{则}\mathcal{O}_{0\cdot}=\begin{bmatrix}0\\ \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}\text{和 }\mathcal{O}_{1\cdot}=\begin{bmatrix}0\\ -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix},$$

那么由$\mathcal{O}_{0\cdot}$和$\mathcal{O}_{1\cdot}$张成的 $\mathbf{R}^3$ 的子空间包含了 $\mathbf{R}^3$ 中的第 $j=0$ 个元素恒等于 0 的所有向量的集合. 除了一开始给出的$\mathcal{O}_{0\cdot}$和$\mathcal{O}_{1\cdot}$，这个子空间显然还存在其他的基底：对于任何 θ，向量

$$\begin{bmatrix}0\\ \cos\theta\\ \sin\theta\end{bmatrix}\quad \text{和}\quad \begin{bmatrix}0\\ -\sin\theta\\ \cos\theta\end{bmatrix}$$

构成了由$\mathcal{O}_{0\cdot}$和$\mathcal{O}_{1\cdot}$张成的子空间的一个基底. 借助于一个术语，我们可以称上面的$\hat{\mathbf{X}}$是 $\mathbf{X}$ 在$\mathcal{O}_{0\cdot}$，…，$\mathcal{O}_{N'-1\cdot}$张成的子空间上的投影.

3.3 复值变换

将上面讨论的内容稍微推广一下，即允许$\mathcal{O}$的元素是复值的，这是有趣的(这将在 3.4 节定义规范正交离散傅里叶变换中用到). 这时，规范正交性质就变成了$\mathcal{O}^H\mathcal{O}=I_N$，其中上标 H 指的是埃尔米特转置，由定义，一个矩阵的埃尔米特转置$\mathcal{O}^H$ 可以这样得到：首先进行普通的转置$\mathcal{O}^T$，然后将它的每一元素用其复共轭代替. 例如，

$$\text{如果}\mathcal{O}=\begin{bmatrix}O_{1,1} & O_{1,2}\\ O_{2,1} & O_{2,2}\end{bmatrix},\quad \text{那么}\mathcal{O}^H=\begin{bmatrix}O_{1,1}^* & O_{2,1}^*\\ O_{1,2}^* & O_{2,2}^*\end{bmatrix},$$

这里$O_{j,k}^*$表示$O_{j,k}$的复共轭. 注意，如果$\mathcal{O}$的元素是实值的，那么$\mathcal{O}^H$ 就等于$\mathcal{O}^T$，因为实变量的复共轭就是它本身. 满足$\mathcal{O}^H\mathcal{O}=I_N$ 的矩阵有时也叫做酉矩阵. 令$\mathcal{O}_{j\cdot}^H$和$\mathcal{O}_{\cdot k}$分别表示$\mathcal{O}$的第 j 个行向量和第 k 个列向量(注意，这些定义与实值情形是一致的，因为在实值情形下$\mathcal{O}_{j\cdot}^H=\mathcal{O}_{j\cdot}^T$). 因此可以写为如下形式：

$$\mathcal{O}=\begin{bmatrix}\mathcal{O}_{0\cdot}^H\\ \mathcal{O}_{1\cdot}^H\\ \vdots\\ \mathcal{O}_{N-1\cdot}^H\end{bmatrix}=[\mathcal{O}_{\cdot 0},\mathcal{O}_{\cdot 1},\cdots,\mathcal{O}_{\cdot N-1}].$$

两个复值向量 $\mathbf{U}$ 和 $\mathbf{V}$ 的内积定义如下：

$$\langle\mathbf{U},\mathbf{V}\rangle=\sum_{t=0}^{N-1}U_tV_t^*=\mathbf{V}^H\mathbf{U}. \tag{45}$$

注意到$\langle \boldsymbol{V}, \boldsymbol{U}\rangle = (\langle \boldsymbol{U}, \boldsymbol{V}\rangle)^*$，那么$\boldsymbol{U}$的范数平方变成

$$\|\boldsymbol{U}\|^2 = \langle \boldsymbol{U},\boldsymbol{U}\rangle = \sum_{t=0}^{N-1} U_t U_t^* = \sum_{t=0}^{N-1} |U_t|^2.$$

根据内积的定义，规范正交性质可以重新表述为：$\langle \mathcal{O}_{\cdot k}, \mathcal{O}_{\cdot k'}\rangle = \delta_{k,k'}$或$\langle \mathcal{O}_{j\cdot}, \mathcal{O}_{j'\cdot}\rangle = \delta_{j,j'}$. $\mathcal{O}_{j\cdot}$和$\mathcal{O}_{\cdot k}$都组成了N维复值向量空间$\mathbf{C}^N$的一个基.

由于实值向量$\boldsymbol{X}$包含在$\mathbf{C}^N$中，所以我们可以把它关于酉矩阵$\mathcal{O}$分解. 当我们关于实值规范正交矩阵分解$\boldsymbol{X}$时，用$\langle \boldsymbol{X}, \mathcal{O}_{j\cdot}\rangle$作为第$j$个变换系数$O_j$的定义(参见式(43a)). 对于酉矩阵，我们同样令$O_j \equiv \langle \boldsymbol{X}, \mathcal{O}_{j\cdot}\rangle$.

练习[46a] 证明

$$\boldsymbol{O} = \mathcal{O}\boldsymbol{X} \quad 和 \quad \boldsymbol{X} = \mathcal{O}^H\boldsymbol{O}, \tag{46a}$$

它们分别定义了基于酉矩阵变换的分解和重构等式(如果我们把$\boldsymbol{X}$换成复值向量，这些等式仍然成立). ◁

注意：对于实值情形，这些等式与式(43a)和(43b)是一致的.

最后，我们注意到能量守恒可以表述为

$$\|\boldsymbol{X}\|^2 = \sum_{t=0}^{N-1} X_t^2 = \sum_{j=0}^{N-1} |O_j|^2 = \|\boldsymbol{O}\|^2,$$

而且投影定理可以很容易地推广到复值变换.

3.4 规范正交离散傅里叶变换

令$\{F_k\}$是实值序列$\{X_t\}$的规范正交离散傅里叶变换，定义为

$$F_k \equiv \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} X_t \mathrm{e}^{-\mathrm{i}2\pi tk/N}, \quad k = 0,\cdots,N-1 \tag{46b}$$

(规范正交离散傅里叶变换和第2章中的离散傅里叶变换的关系将在本节最后的评论与扩展中的条目[1]中讨论). 量F_k称为第k个傅里叶系数. 虽然F_k是N个复值变量，但是组成它们实部和虚部的$2N$个实数值是有约束的，即只有N个实数值是构造所有傅里叶系数所需要的.

练习[46b] 令$\{X_t: t=0, 1, \cdots, N-1\}$是一个实值时间序列，$F_k$是式(46b)定义的第$k$个傅里叶系数. 证明$F_0$是实值的，$F_{N-k} = F_k^*$对于$1 \leqslant k < N/2$成立，并且当$N$为偶数时，$F_{N/2}$是实值的. ◁

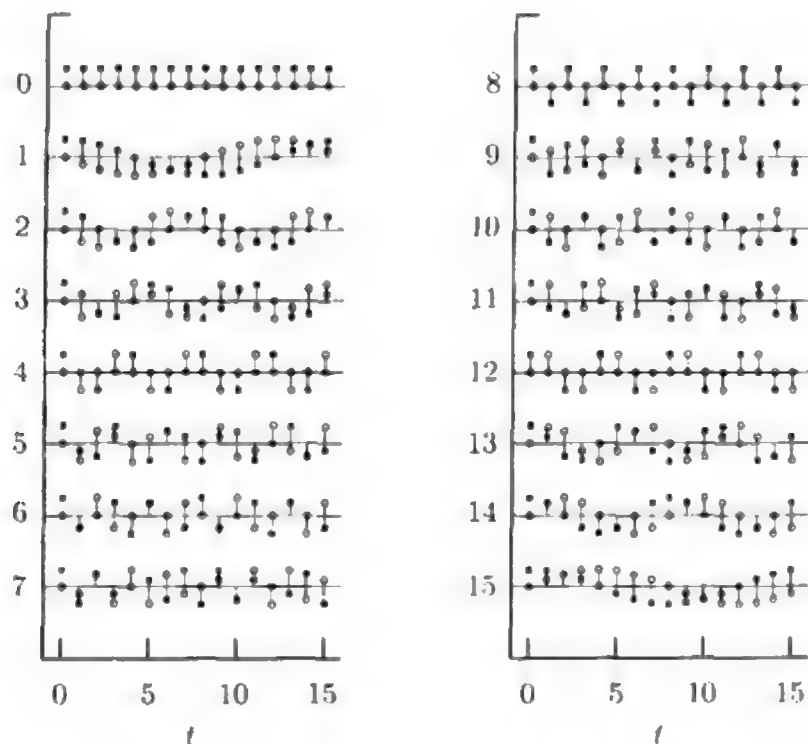


图 47 $N=16$ 且 $k=0$ 到 7(左边插图从上到下)和 $k=8$ 到 15(右边的插图)的规范正交离散傅里叶变换矩阵 $\mathcal{F}$ 的行向量 $\mathcal{F}_{k\cdot}^H$. $\mathcal{F}$ 的大多数元素是复值的，因此实部和虚部分别由实方块和空心圆表示. 注意 $\mathcal{F}_{0\cdot}^H$ 和 $\mathcal{F}_{8\cdot}^H$ 是实值的，并且，例如 $\mathcal{F}_{15\cdot}^H=(\mathcal{F}_{1\cdot}^H)^*$

从上面的练习我们知道 F_k 由 N 个实数决定，即由 F_0，$F_{N/2}$(当 N 是偶数时)以及复值变量 F_k，$1\leqslant k<N/2$ (当 N 是偶数时有 $N/2-1$ 个，当 N 为奇数时有$(N-1)/2$ 个)的实部和虚部分量所决定.

利用向量的记号我们可以把式(46b)写为 $\boldsymbol{F}=\mathcal{F}\,\boldsymbol{X}$，这里 $\boldsymbol{F}$ 是长度为 N 的列向量，它的元素是 F_0，F_1，…，F_{N-1}，$\mathcal{F}$ 是一个 $N\times N$ 的矩阵，它的第(k,t)个元素是 $\exp(-\mathrm{i}2\pi tk/N)/\sqrt{N}$，$0\leqslant k$，$t\leqslant N-1$，图 47 示例了一个 $N=16$ 时的矩阵 $\mathcal{F}$.

练习[47] 验证 $\mathcal{F}$ 是酉矩阵，即 $\mathcal{F}^H\,\mathcal{F}=I_N$，这里上标 H 代表埃尔米特转置.

提示：利用练习[29a]的提示. ◁

由上述可以得到 $\boldsymbol{X}=\mathcal{F}^H\boldsymbol{F}$，并且

$$\varepsilon_F\equiv\|\boldsymbol{F}\|^2=\|\boldsymbol{X}\|^2\equiv\varepsilon_X.$$

上式表明规范正交离散傅里叶变换是保持能量的(等距)变换. 由于 $F_{N-k}=F_k^*$，当 N 为偶数时，有

$$\|\boldsymbol{F}\|^2=\sum_{k=0}^{N-1}|F_k|^2=|F_0|^2+|F_{N/2}|^2+2\sum_{1\leqslant k<N/2}|F_k|^2.$$

这样，我们就可以将$\{X_t\}$的能量关于指标$k(0\leqslant k\leqslant N/2)$进行分解(如果$N$是奇数，就去掉$F_{N/2}$项)．规范正交离散傅里叶变换在很多应用方面被认为有用的，因为傅里叶系数F_k和频率$f_k\equiv k/N$能联系起来，因此规范正交离散傅里叶变换可以认为是关于"频率-频率"基进行能量分解．建立这种关系的一种途径就是注意到F_k的计算涉及$\exp(-\mathrm{i}2\pi tk/N)=\exp(-\mathrm{i}2\pi tf_k)$，利用欧拉(Euler)公式$\exp(-\mathrm{i}2\pi tf_k)=\cos(2\pi tf_k)-\mathrm{i}\sin(2\pi tf_k)$，就可以将$f_k$和正弦频率联系起来．另一种方法是，如果对于固定的振幅A、固定的频率$f_{k'}(0\leqslant k'\leqslant N/2)$和固定的相位$\phi$，有$X_t=A\cos(2\pi f_{k'}t+\phi)$，那么对于$0\leqslant k\leqslant N/2$，我们有

$$|F_k|^2=\begin{cases}NA^2/4, & k=k'\neq 0 \text{ 或 } N/2;\\ NA^2\cos^2(\phi), & k=k'=0 \text{ 或 } N/2;\\ 0, & \text{其他}.\end{cases} \tag{48a}$$

即当指标从0到$N/2$时，频率为$f_{k'}$的正弦的规范正交离散傅里叶变换仅有一个单独的非零系数，并且这个系数和频率$f_{k'}$有关(练习[3.4]验证了这个结果)．

在一些特殊应用中，我们必须考虑每次观察的时间间隔Δt，这时标准的频率k/N对应于物理频率$k/(N\Delta t)$．例如，如果$\Delta t=1$秒，那么$k/(N\Delta t)$就是频率，每秒k/N次周期(即赫兹，通常缩写为Hz)．

令

$$\overline{X}\equiv\frac{1}{N}\sum_{t=0}^{N-1}X_t$$

是X_t的采样平均，并且令

$$\hat{\sigma}_X^2\equiv\frac{1}{N}\sum_{t=0}^{N-1}(X_t-\overline{X})^2=\frac{1}{N}\sum_{t=0}^{N-1}X_t^2-\overline{X}^2=\frac{1}{N}\|\mathbf{X}\|^2-\overline{X}^2 \tag{48b}$$

是采样方差．因为$F_0/\sqrt{N}=\overline{X}$，对于偶数$N$采样方差可以分解成

$$\hat{\sigma}_X^2=\frac{1}{N}\|\boldsymbol{F}\|^2-\overline{X}^2=\frac{1}{N}|F_{N/2}|^2+\frac{2}{N}\sum_{1\leqslant k<N/2}|F_k|^2,$$

因此$2|F_k|^2/N$表示对频率为$f_k(0<k<N/2)$的$\{X_t\}$的采样方差贡献(对于奇数N，我们必须去掉$F_{N/2}$项)．对于偶数N，方差的分解可以用来定义$\{X_t\}$的离散傅里叶经验(实验)功率(即取样变量)谱$\{P_{\mathcal{F}}(f_k):f_k\equiv k/N,\ k=1,\cdots,N/2\}$

$$P_{\mathcal{F}}(f_k)\equiv\begin{cases}2|F_k|^2/N, & 1\leqslant k<N/2;\\ |F_k|^2/N, & k=N/2.\end{cases} \tag{48c}$$

根据这个定义，我们有$\sum_{k=1}^{N/2}P_{\mathcal{F}}(f_k)=\hat{\sigma}_X^2$(对于奇数$N$,我们还是去掉$F_{N/2}$项,功率谱是定义于频率$f_k=k/N,k=1,2,\cdots,(N-1)/2$)．对于奇数或者偶数$N$，$P_{\mathcal{F}}(f_k)$

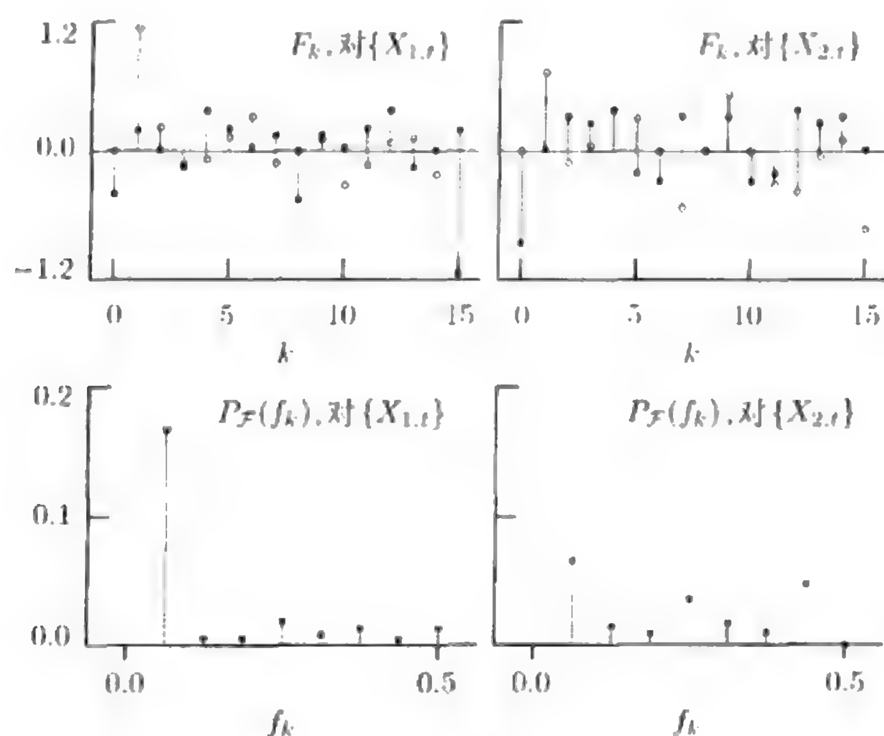


图 49 图 42 所示时间序列的规范正交离散傅里叶变换. $\{X_{1,t}\}$ 和 $\{X_{2,t}\}$ 的规范正交离散傅里叶变换 $\boldsymbol{F}$ 如上面一行所示. 规范正交离散傅里叶变换的实部和虚部分别由实的方块和空心圆表示, 对应于规范正交离散傅里叶变换的离散经验功率谱在下面一行给出

中的频率 f_k 满足 $1/N \leqslant f_k \leqslant 1/2$.

作为例子, 图 49 最上边一行所示的是图 42 中两个 16 点时间序列 $\{X_{1,t}\}$ 和 $\{X_{2,t}\}$ 的规范正交离散傅里叶变换. 对应的离散经验功率谱如下面一行所示. 尽管这两个时间序列仅在第 13 个值不同, 但是他们的规范正交离散傅里叶变换和谱却在每一个点都不同.

现在我们考虑由式(46a)表示的 $\boldsymbol{X}$ 的傅里叶重构

$$\boldsymbol{X} = \mathcal{F}^H \boldsymbol{F} = \sum_{k=0}^{N-1} \langle \boldsymbol{X}, \mathcal{F}_{k\cdot} \rangle \mathcal{F}_{k\cdot} = \sum_{k=0}^{N-1} F_k \mathcal{F}_{k\cdot}.$$

考虑这个向量式的第 t 个元素, 我们得到相应式(46b)规范正交离散傅里叶变换的逆

$$X_t = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} F_k e^{i2\pi tk/N}, \quad t = 0, \cdots, N-1.$$

下面, 我们注意到, 对于偶数 N, 记

$$\sum_{k=0}^{N-1} F_k \mathcal{F}_{k\cdot} = F_0 \mathcal{F}_{0\cdot} + \sum_{1 \leqslant k < N/2} (F_k \mathcal{F}_{k\cdot} + F_{N-k} \mathcal{F}_{N-k\cdot}) + F_{N/2} \mathcal{F}_{N/2\cdot}$$

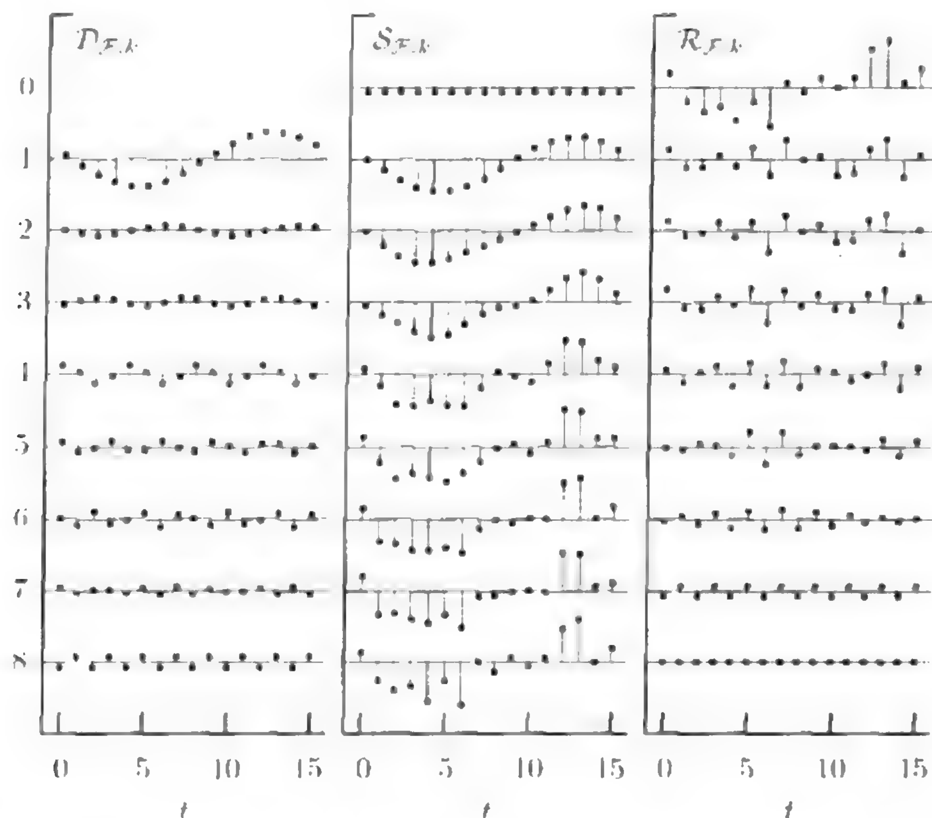


图50 $\{X_{1,t}\}$(图42左边一行所示)的傅里叶细节$\mathcal{D}_{\mathcal{F},k}$、光滑$\mathcal{S}_{\mathcal{F},k}$和粗糙$\mathcal{R}_{\mathcal{F},k}$，$k=0,\cdots,8$(从上到下)

(对于奇数我们同样去掉$F_{N/2}$项). 让我们定义

$$\mathcal{D}_{\mathcal{F},k} \equiv \begin{cases} F_k \mathcal{F}_{k\cdot} + F_{N-k} \mathcal{F}_{N-k\cdot}, & 1 \leqslant k < N/2; \\ F_{N/2} \mathcal{F}_{N/2\cdot}, & k = N/2. \end{cases} \tag{50}$$

练习[50a] 证明$\mathcal{D}_{\mathcal{F},k} = 2\,\Re(F_k \mathcal{F}_{k\cdot})$，$0<k<N/2$，其中$\Re(F_k \mathcal{F}_{k\cdot})$是一个向量，它的第$k$个元素是复变量$F_k$和向量$\mathcal{F}_{k\cdot}$的第$k$个元素的乘积的实部. 提示：先证明$F_{N-k}\mathcal{F}_{N-k\cdot} = F_k^* \mathcal{F}_{k\cdot}^*$. ◁

由于$F_0 \mathcal{F}_{0\cdot} = \overline{X}\mathbf{1}$，其中$\mathbf{1}$是一个由1构成的向量，则

$$\mathbf{X} = \overline{X}\mathbf{1} + \sum_{k=1}^{\lfloor N/2 \rfloor} \mathcal{D}_{\mathcal{F},k},$$

其中$\lfloor N/2 \rfloor$表示不超过$N/2$的最大整数(因此，如果N是偶数，则$\lfloor N/2 \rfloor = N/2$；如果$N$是奇数，则$\lfloor N/2 \rfloor = (N-1)/2$). 上面的公式根据实值向量$\mathcal{D}_{\mathcal{F},k}$的分解提供了一个额外的$\mathbf{X}$，其中每一个$\mathcal{D}_{\mathcal{F},k}$都和一个频率$f_k$相联系.

练习[50b] 令$F_k \equiv A_k - \mathrm{i}B_k$，并令$\mathcal{D}_{\mathcal{F},k,t}$为$\mathcal{D}_{\mathcal{F},k}$($1\leqslant k<N/2$)的第$t$个元素，证明

$$\mathcal{D}_{\mathcal{F},k,t} = \frac{2}{\sqrt{N}}[A_k \cos(2\pi f_k t) + B_k \sin(2\pi f_k t)],$$

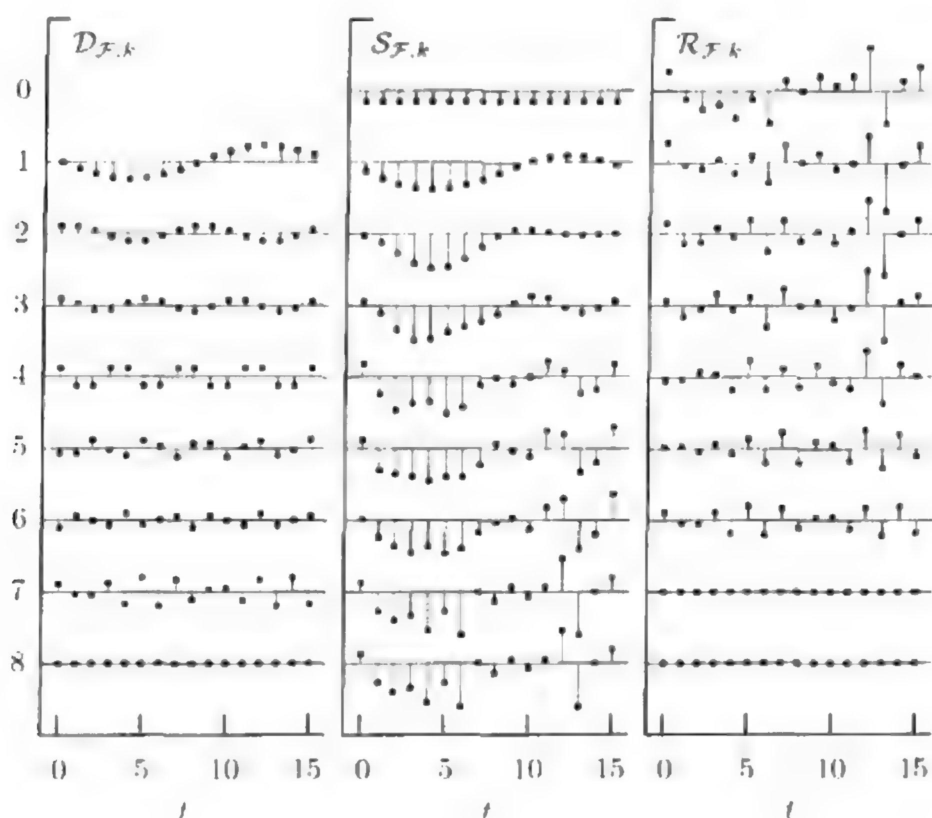


图 51 $\{X_{2,t}\}$(图 42 右边一行所示)的傅里叶细节$\mathcal{D}_{\mathcal{F},k}$、光滑$\mathcal{S}_{\mathcal{F},k}$和粗糙$\mathcal{R}_{\mathcal{F},k}$，$k=0$，…，8(从上到下)

它表明$\mathcal{D}_{\mathcal{F},k}$是实的并且和频率 f_k 有关. ◁

我们称$\mathcal{D}_{\mathcal{F},k}$为 k 阶傅里叶细节. 图 50 和 51 的第一列分别显示了图 42 中的两个时间序列$\{X_{1,t}\}$和$\{X_{2,t}\}$的傅里叶细节(这里$\lfloor N/2\rfloor=8$). 注意到虽然这两个时间序列仅在 $t=13$ 处的值不一样，但是它们的$\mathcal{D}_{\mathcal{F},k}$却在所有的 k 处都不相同.

令

$$\mathcal{S}_{\mathcal{F},k}\equiv\overline{X}\mathbf{1}+\sum_{j=1}^{k}\mathcal{D}_{\mathcal{F},j},\quad \mathcal{R}_{\mathcal{F},k}\equiv\sum_{j=k+1}^{\lfloor N/2\rfloor}\mathcal{D}_{\mathcal{F},j},$$

因此，我们有 $\mathbf{X}=\mathcal{S}_{\mathcal{F},k}+\mathcal{R}_{\mathcal{F},k}$，$0\leqslant k\leqslant\lfloor N/2\rfloor$(这里$\mathcal{S}_{\mathcal{F},0}\equiv\overline{X}\mathbf{1}$，并且定义$\mathcal{R}_{\mathcal{F},\lfloor N/2\rfloor}$是一个零向量). 因为$\mathcal{S}_{\mathcal{F},k}$是由采样平均和 $\mathbf{X}$ 的傅里叶重构中的 k 个最低频成分构成的，而$\mathcal{R}_{\mathcal{F},k}$由$\lfloor N/2\rfloor-k$ 个高频成分组成，因此由$\mathcal{S}_{\mathcal{F},k}$的元素给出的时间序列应该看起来很光滑，相反的$\mathcal{R}_{\mathcal{F},k}$看起来很粗糙(术语“光滑”和“粗糙”是由 Tukey 在 1977 年引入的). 我们称$\mathcal{S}_{\mathcal{F},k}$和$\mathcal{R}_{\mathcal{F},k}$分别为 $\mathbf{X}$ 的 k 阶傅里叶光滑和傅里叶粗糙(工程师们通常称他们分别为 $\mathbf{X}$ 的低通和高通). 图 50 和图 51 的第 2 和第 3 列分别示例了时间序列$\{X_{1,t}\}$和$\{X_{2,t}\}$的所有阶的傅里叶光滑和粗糙.

注意到我们可以用投影定理把$\mathcal{S}_{\mathcal{F},k}$解释成$\boldsymbol{X}$基于基向量$\mathcal{F}_{0\cdot}$，$\mathcal{F}_{1\cdot}$，…，$\mathcal{F}_{k\cdot}$和$\mathcal{F}_{N-k\cdot}$，$\mathcal{F}_{N-k+1\cdot}$，…，$\mathcal{F}_{N-1\cdot}$(最小二乘意义下)的最佳逼近(如果$k=0$，那么第二个集合就是空的)．相似地，$\mathcal{R}_{\mathcal{F},k}$，$0\leqslant k<\lfloor N/2\rfloor$是$\boldsymbol{X}$基于$\mathcal{F}_{k+1\cdot}$，$\mathcal{F}_{k+2\cdot}$，…，$\mathcal{F}_{N-k-1\cdot}$的最佳逼近．

最后，我们考虑向量$\boldsymbol{X}$的循环移位对规范正交离散傅里叶变换的影响．设$\mathcal{T}$和$\mathcal{T}^{-1}$是$N\times N$阶矩阵，满足

$$\mathcal{T}\boldsymbol{X}\equiv[X_{N-1},X_0,X_1,\cdots,X_{N-3},X_{N-2}]^{\mathrm{T}} \tag{52a}$$

和

$$\mathcal{T}^{-1}\boldsymbol{X}\equiv[X_1,X_2,X_3,\cdots,X_{N-2},X_{N-1},X_0]^{\mathrm{T}},$$

并且令$\mathcal{T}^2\boldsymbol{X}\equiv\mathcal{T}\mathcal{T}\boldsymbol{X}$，$\mathcal{T}^{-2}\boldsymbol{X}\equiv\mathcal{T}^{-1}\mathcal{T}^{-1}\boldsymbol{X}$等．注意到对于所有的$\boldsymbol{X}$有$\mathcal{T}^{-1}\mathcal{T}\boldsymbol{X}=\boldsymbol{X}$，因此$\mathcal{T}^{-1}\mathcal{T}=I_N$，故$\mathcal{T}$的逆是$\mathcal{T}^{-1}$(正如符号所暗示的)．$\mathcal{T}$和$\mathcal{T}^{-1}$所取的确切形式分别为

$$\begin{bmatrix}0&0&0&0&\cdots&0&0&1\\1&0&0&0&\cdots&0&0&0\\0&1&0&0&\cdots&0&0&0\\\vdots&\vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\\0&0&0&0&\cdots&1&0&0\\0&0&0&0&\cdots&0&1&0\end{bmatrix}\quad\text{和}\quad\begin{bmatrix}0&1&0&0&\cdots&0&0&0\\0&0&1&0&\cdots&0&0&0\\0&0&0&1&\cdots&0&0&0\\\vdots&\vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\\0&0&0&0&\cdots&0&0&1\\1&0&0&0&\cdots&0&0&0\end{bmatrix} \tag{52b}$$

从这可以明显地看出$\mathcal{T}^{-1}=\mathcal{T}^{\mathrm{T}}$，即$\mathcal{T}$实际上是一个规范正交变换．

练习[52] 对于任意的整数m，证明$\mathcal{T}^m\boldsymbol{X}$的第$k$个傅里叶系数是$F_k\exp(-\mathrm{i}2\pi mk/N)$． ◁

由于$|F_k\exp(-\mathrm{i}2\pi mk/N)|=|F_k|$，因此离散傅里叶变换的经验功率谱$\{P_{\mathcal{F}}(f_k)\}$关于$\boldsymbol{X}$的循环移位是不变的．

3.4 节的评论与扩展

[1]对式(46b)的规范正交离散傅里叶变换和式(36g)的离散傅里叶变换进行对比可以发现，这两个变换仅仅差了一个因子$1/\sqrt{N}$．因此，如果我们构造了一个$N\times N$维的矩阵$\bar{\mathcal{F}}$，它的第(k,t)个元素是$\exp(-\mathrm{i}2\pi tk/N)$，那么$\bar{\mathcal{F}}\boldsymbol{X}$将生成一个包含$\boldsymbol{X}$的离散傅里叶变换的向量．由于$\bar{\mathcal{F}}^H\bar{\mathcal{F}}$不等于单位矩阵$I_N$而等于$N$乘以$I_N$，因此可以看出离散傅里叶变换不是规范正交变换，但是可以叫做正交变换，因为$\bar{\mathcal{F}}^H\bar{\mathcal{F}}$的所有非对角线上的元素都是零．正如我们所定义的那样，采用离散傅里叶变换而非正交离散傅里叶变换的理性来自于卷积的离散傅里叶变换等于每一个离散傅里叶变换的乘积，然而卷积的正交离散傅里叶变换等于每个正交离散傅里叶变换的乘积的$\sqrt{N}$倍．卷积在小波理论以及其他地方的应用非常广泛，没有显示的因子$\sqrt{N}$是非常方便的．

[2]规范正交离散傅里叶变换固有的周期性假设对于许多时间序列是不合适的，特别是那些$|X_0-X_{N-1}|$比相邻观察的典型绝对差(即$|X_t-X_{t-1}|$，$t=2,\cdots,N$)大很多的那些时间序列．已经发现两个很有用的程序，可以用来

减弱对产生的分析的影响，即逐渐变细(有时称为加窗)和将 $\boldsymbol{X}$ 延拓一倍，通过在 $\boldsymbol{X}$ 的结尾加上它的一个倒序形式(第二个程序将在 4.11 节讨论). 逐渐变细就是给一个序列$\{a_t\}$乘上 $\boldsymbol{X}$ 的元素，a_t 当 t 趋于 0 或者 $N-1$ 时减小到 0，使得 $\boldsymbol{X}$ 的中间部分的形状基本上没有太大的改变. 通过强制时间序列在它的终点处趋于 0，我们发现 $a_0X_0 \approx a_{N-1}X_{N-1}$，因此周期性的假设是有些合理的.

正如我们将在 4.11 节看到的那样，离散小波变换也假设周期性，但是变换的局部化性质使得受影响的系数的个数比规范正交离散傅里叶变换情形要小得多.

3.5 小结

这里我们总结一下本章的主要观点. 一个 $N\times N$ 阶实值矩阵$\mathcal{O}$是规范正交的，如果它的转置是它的逆，即如果$\mathcal{O}^T\mathcal{O}=I_N$，其中 I_N 是 $N\times N$ 阶单位矩阵. 当然还有$\mathcal{O}\mathcal{O}^T=I_N$. 我们可以用$\mathcal{O}$左乘 $\boldsymbol{X}$ 来分析(分解)时间序列 $\boldsymbol{X}$ 得到 $\boldsymbol{O}=\mathcal{O}\boldsymbol{X}$，一个包含变换系数的 N 维列向量；同样，给定 $\boldsymbol{O}$，我们可以合成(重构)$\boldsymbol{X}$ 因为$\mathcal{O}^T\boldsymbol{O}=\boldsymbol{X}$. 如果令$\mathcal{O}_{j\cdot}$是一个列向量，它的元素包含了$\mathcal{O}$的第 j 行，那么 $\boldsymbol{O}$ 的第 j 个元素 O_j 由内积$\langle \boldsymbol{X}, \mathcal{O}_{j\cdot}\rangle$给出，因此我们可以将 $\boldsymbol{X}$ 的重构表示成

$$\boldsymbol{X}=\mathcal{O}^T\boldsymbol{O}=\sum_{j=0}^{N-1}O_j\mathcal{O}_{j\cdot}=\sum_{j=0}^{N-1}\langle \boldsymbol{X},\mathcal{O}_{j\cdot}\rangle\,\mathcal{O}_{j\cdot}.$$

一个规范正交变换是等距的，这意味着在 $\sum_t X_t^2=\|\boldsymbol{X}\|^2=\|\boldsymbol{O}\|^2$ 的意义下是保持能量的. 投影定理表明，$\boldsymbol{X}$ 的最佳逼近$\hat{\boldsymbol{X}}$仅由那些 $N'<N$ 的向量$\mathcal{O}_{0\cdot},\cdots,\mathcal{O}_{N'-1}$构成

$$\hat{\boldsymbol{X}}=\sum_{j=0}^{N'-1}O_j\mathcal{O}_{j\cdot},$$

这里的“最佳”是在最小二乘意义下理解的，即误差向量 $\boldsymbol{e}\equiv\boldsymbol{X}-\hat{\boldsymbol{X}}$的范数最小.

对于元素为复值的矩阵$\mathcal{O}$也有类似的结果. 如果$\mathcal{O}^H\mathcal{O}=I_N$，其中上标 H 表示埃尔米特转置，那么这个矩阵就叫做酉矩阵. 当然，$\mathcal{O}\mathcal{O}^H=I_N$. 如上，实值向量 $\boldsymbol{X}$ 的变换系数由$\boldsymbol{O}=\mathcal{O}\boldsymbol{X}$ 给出，其第 j 个元素是$O_j=\langle \boldsymbol{X}, \mathcal{O}_{j\cdot}\rangle$，其中$\mathcal{O}_{j\cdot}$是$\mathcal{O}$的第 j 行的埃尔米特转置. 重构公式由下式得到

$$\boldsymbol{X}=\mathcal{O}^H\boldsymbol{O}=\sum_{j=0}^{N-1}O_j\mathcal{O}_{j\cdot}=\sum_{j=0}^{N-1}\langle \boldsymbol{X},\mathcal{O}_{j\cdot}\rangle\,\mathcal{O}_{j\cdot}.$$

变换$\mathcal{O}$保持能量：$\|\boldsymbol{X}\|^2=\|\boldsymbol{O}\|^2=\sum_t|O_j|^2$. 当 $\boldsymbol{X}$ 用复值向量代替时，上面的结果仍然成立.

前面的复值规范正交变换的例子就是规范正交离散傅里叶变换(ODFT)，是由 $N\times N$ 矩阵$\mathcal{F}$定义的，它的第(k,t)个元素是 $\exp(-i2\pi tk/N)/\sqrt{N}$(规范正交

离散傅里叶变换和离散傅里叶变换的区别就在于因子，参看式(46b)和(36g)). 这个变换的每一个系数都和一个特定的频率，即 $f_k=k/N$，联系，容许对能量按频率分解. 通过按频率对重构向量进行分组，我们就可以将 $\boldsymbol{X}$ 分解成光滑部分 $\mathcal{S}_{\mathcal{F},k}$($\boldsymbol{X}$ 的低通部分)和粗糙部分 $\mathcal{R}_{\mathcal{F},k}$($\boldsymbol{X}$ 的高通部分)，因此有 $\boldsymbol{X}=\mathcal{S}_{\mathcal{F},k}+\mathcal{R}_{\mathcal{F},k}$. 细节 $\mathcal{D}_{\mathcal{F},k}$ 和频率 f_k 有关，它表示相邻两个顺序的光滑部分(或者粗糙部分)的差分，即 $\mathcal{D}_{\mathcal{F},k}=\mathcal{S}_{\mathcal{F},k}-\mathcal{S}_{\mathcal{F},k-1}$(或 $\mathcal{D}_{\mathcal{F},k}=\mathcal{R}_{\mathcal{F},k-1}-\mathcal{R}_{\mathcal{F},k}$).

3.6 练习

[3.1] 证明规范正交变换 $\mathcal{O}$ 保持内积，即

$$\langle \mathcal{O}\boldsymbol{X},\mathcal{O}\boldsymbol{Y}\rangle=\langle \boldsymbol{X},\boldsymbol{Y}\rangle.$$

[3.2] 这里我们定义一个变换，它和规范正交离散傅里叶变换非常类似，不同之处仅在于这个变换直接用正弦和余弦而不是借助于复指数. 设 N 是一个偶数，假设 $\mathcal{O}$ 是一个 $N\times N$ 阶矩阵，它的第 (k,t) 个元素是 $\mathcal{O}_{k,t}$，$0\leqslant k,t\leqslant N-1$，变换由下式给出：

$$\mathcal{O}_{k,t}=\begin{cases}\sqrt{2/N}\cos([k+2]\pi t/N), & k=0,2,\cdots,N-4;\\ \sqrt{2/N}\sin([k+1]\pi t/N), & k=1,3,\cdots,N-3;\\ \cos(\pi t)/\sqrt{N}, & k=N-2;\\ 1/\sqrt{N}, & k=N-1.\end{cases}$$

证明 $\mathcal{O}$ 是规范正交矩阵. 对于奇数 N 我们如何定义呢?

[3.3] 这里我们考虑类型Ⅱ和类型Ⅳ的离散余弦变换(DCT)，它们都是规范正交变换，工程师们对此应用很广泛，例如在图像压缩中，详见 Rao and Yip(1990).

(a)设 $\mathcal{O}$ 是一个 $N\times N$ 阶矩阵，第 (k,t) 个元素是

$$\frac{\sqrt{2}}{\sqrt{c_k N}}\cos\left(\pi k\frac{t+0.5}{N}\right),\quad k,t=0,\cdots,N-1,$$

其中 $c_0=2$，$c_k=1$，$k=1,\cdots,N-1$(注意 $k=$第 0 行的元素都是 $1/\sqrt{N}$). 证明 $\mathcal{O}$ 是一个规范正交矩阵. 这个变换就是人们熟知的第Ⅱ型离散余弦变换.

(b)设 $\mathcal{O}$ 是一个 $N\times N$ 阶矩阵，第 (k,t) 个元素是

$$\frac{\sqrt{2}}{\sqrt{N}}\cos\left(\pi(k+0.5)\frac{t+0.5}{N}\right),\quad k,t=0,\cdots,N-1.$$

证明 $\mathcal{O}$ 是一个规范正交矩阵. 这个变换是第Ⅳ型离散余弦变换.

(与正交离散傅里叶变换和第Ⅱ型离散余弦变换不同，第Ⅳ型离散余弦变换总是实值的；和正交离散傅里叶变换类似，每一个离散余弦变换系数可以和一个频率联系，但是要注意到每个相邻频率的间隔是 $1/2N$，是离散傅里叶变换频率间隔即 $1/N$ 两倍精度. 和正交离散傅里叶变换类似，第Ⅱ型离散余弦变换的基向量中有一个是常数，因此

第Ⅱ型离散余弦变换系数中有一个系数和采样平均 $\bar{X}$ 成正比；这对于第Ⅳ型离散余弦变换不成立，而它对于采样平均为 0 的序列是非常有用的.）

[3.4] 如式(46b)所定义，令 F_k 是长度为 N 的序列 X_0，…，X_{N-1} 的规范正交离散傅里叶变换的第 k 个系数. 假定 $X_t = A\cos(2\pi f_{k'} t + \phi)$，其中 A 是固定的振幅，$f_{k'} = k'/N$ 是固定的频率，$0 \leqslant k' \leqslant N/2$，$\phi$ 是固定的相位. 证明 $|F_k|^2$ 由式(48a)给出.

[3.5] 证明：对于 $f_k \equiv k/N$，$1 \leqslant k \leqslant N/2$，$P_{\mathcal{F}}(f_k) = \|\mathcal{D}_{\mathcal{F},k}\|^2/N$，其中 $P_{\mathcal{F}}(f_k)$ 是离散傅里叶经验功率谱(在式(48c)中定义)的第 k 个分量，而 $\mathcal{D}_{\mathcal{F},k}$ 是第 k 阶傅里叶细节(在式(50)中定义).

[3.6] 证明，如果 $\mathcal{D}_{\mathcal{F},k}$ 是 $\boldsymbol{X}$ 的 k 阶傅里叶细节，那么 $\mathcal{T}^m \mathcal{D}_{\mathcal{F},k}$ 是 $\mathcal{T}^m \boldsymbol{X}$ 的 k 阶傅里叶细节，其中 $\mathcal{T}$ 在式(52a)中定义，m 是任意的整数.

[3.7] 验证图 50 的内容($\{X_{1,t}\}$ 的值在图 42 的标题中给出).

[3.8] 确定循环滤波器 $\{a_{k,t}: t = 0, \cdots, N-1\}$ 使得用 $\{a_{k,t}\}$ 对 $\boldsymbol{X}$ 滤波产生以傅里叶光滑 $\mathcal{S}_{\mathcal{F},k}$ 分量的输出. 这个滤波器可以用离散傅里叶变换来实现吗?

[3.9] 考虑下述形式的规范正交矩阵 $\mathcal{O}$：

$$
\mathcal{O} = \begin{bmatrix}
-\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}
\end{bmatrix}
$$

(我们将在研究部分哈尔离散小波变换的时候遇到这样的矩阵). 令 $\boldsymbol{X}$ 是图 42 所示的包含 16 个值的时间序列 $\{X_{1,t}\}$ 向量(时间序列的值在图的标题中给出). 使用 16×16 型的矩阵 $\mathcal{O}$，计算并且绘出 $\boldsymbol{X}$ 在最小二乘意义下基于 $\mathcal{O}_{0\bullet}$，$\mathcal{O}_{1\bullet}$，…，$\mathcal{O}_{7\bullet}$(即从 $\mathcal{O}$ 的前 8 行的元素所形成的向量)的逼近 $\hat{\boldsymbol{X}}_{0,\cdots,7}$ —见 3.2 节中关于投影定理的讨论. 重复上面的步骤可以得到基于 $\mathcal{O}_{8\bullet}$，$\mathcal{O}_{9\bullet}$，…，$\mathcal{O}_{15\bullet}$ 的最小二乘逼近 $\hat{\boldsymbol{X}}_{8,\cdots,15}$. $\boldsymbol{X}$ 能否从 $\hat{\boldsymbol{X}}_{0,\cdots,7}$ 和 $\hat{\boldsymbol{X}}_{8,\cdots,15}$ 中重构 $\boldsymbol{X}$?

第 4 章

离散小波变换

4.0 引言

在这一章，我们引入离散小波变换(DWT)，它是用小波变换研究时间序列的基本工具，其作用有如离散傅里叶变换在频谱分析中所起的作用．我们只假定读者熟悉第 2、3 章给出的线性滤波理论和线性代数的基本概念．我们将在这些基本概念上逐渐深入，对于很熟悉这些领域的读者，这也许是不必要的．我们鼓励这些读者只读本章每一节的关键结论和定义，或者直接跳到 4.12 节，这一节包含了离散小波变换简要的自含讨论．若想进一步了解离散小波变换，可参看 Strang(1989，1993)，Rioul and Vetterli(1991)，Press 等(1992)和 Mulcahy(1996)．

这一章下面的部分是按如下方式组织的．4.1 节主要用哈尔小波和 D(4)小波作为例子，给出了离散小波变换的一个定性描述．离散小波变换严格的数学推导从 4.2 节开始，给出了小波滤波器的定义，并且讨论了一个滤波器是小波滤波器所必须满足的基本条件．4.3 节给出了尺度滤波器，它可以从小波滤波器以一种简单的方式构造出来．小波滤波器和尺度滤波器可以一起用来确定计算(和精确定义)离散小波变换的塔式算法．在 4.4、4.5 和 4.6 节我们将详细讨论这个算法．4.7 节讨论了“部分”离散小波变换的概念，接着我们在 4.8 和 4.9 节讨论了两类特殊的小波滤波器(分别是 Daubechies 滤波器和 Coiflet 滤波器)．4.10 节将给出离散小波变换分析心电图数据的一个实例．4.11 节讨论了实际中计算离散小波变换中会碰到的一些问题(如何选择小波，边界条件的处理问题，如何处理不是 2 的整数幂长的时间序列，选择“部分”离散小波变换的分解层数等问题)．4.12 节总结了这一章的内容．

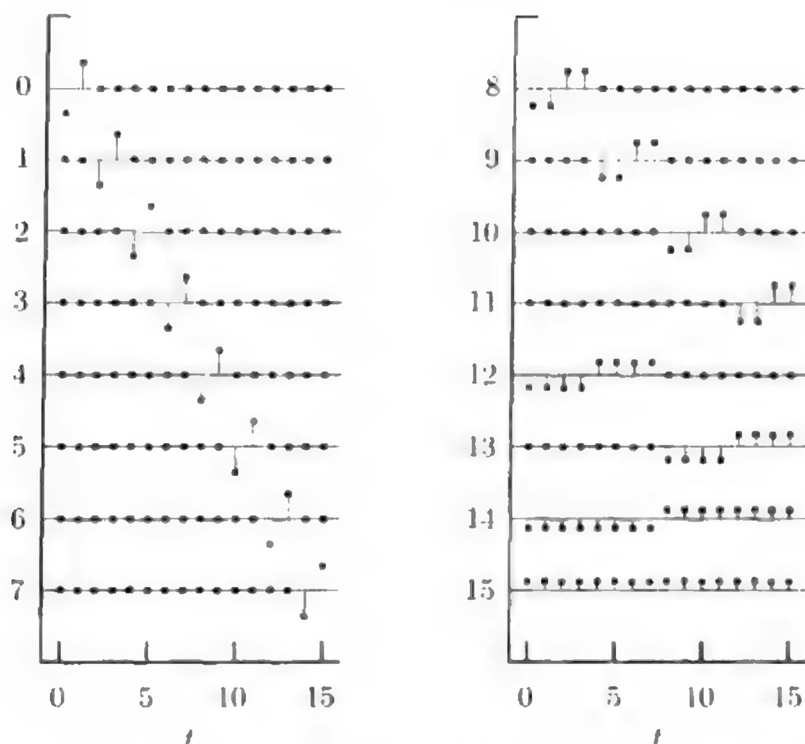


图 57　基于哈尔小波的离散小波变换的矩阵的行向量$\mathcal{W}_{n\cdot}^{\mathrm{T}}$，其中 $N=16$，$n=0$ 到 7(左图从上到下)，$n=8$ 到 15(右图)

4.1　离散小波变换的定性描述

和 3.4 节中讨论的规范正交离散傅里叶变换(ODFT)一样，序列$\{X_t\}$的离散小波变换(DWT)也是一种规范正交变换. 如果我们用$\{W_n: n=0, \cdots, N-1\}$表示离散小波变换的系数，则可以将序列 $\boldsymbol{X}$ 的离散小波变换写成 $\boldsymbol{W}=\mathcal{W}\boldsymbol{X}$，其中 $\boldsymbol{W}$ 是长度为 $N=2^J$ 的列向量，其第 n 个元素是第 n 个离散小波变换系数 W_n. $\mathcal{W}$ 是 $N\times N$ 的实值矩阵，它定义了离散小波变换且满足$\mathcal{W}^{\mathrm{T}}\mathcal{W}=I_N$(需要限制序列 $\boldsymbol{X}$ 的长度为 2^J，参见 4.7、4.11 和 5.0 节的讨论)，类似于规范正交离散傅里叶变换，规范正交性条件意味着 $\boldsymbol{X}=\mathcal{W}^{\mathrm{T}}\boldsymbol{W}$ 且 $\|\boldsymbol{W}\|^2=\|\boldsymbol{X}\|^2$，因此 W_n^2 表示第 n 个离散小波变换系数贡献的能量.

规范正交离散傅里叶变换的系数是和频率相联系的，而和离散小波变换的第 n 个小波系数 W_n 相联系的是某一特定尺度和特定时间段. 为了说明这一点，图 57 给出了对于 $N=16$ 的哈尔离散小波变换的$\mathcal{W}$的元素用一行行的实心方块表示出来，其简单的结构便于我们说明问题(历史上哈尔离散小波变换可以认为是第一个离散小波变换，因为它和哈尔在 1910 年构造的在区间$(-\infty, \infty)$上的平方可积函数空间的规范正交基是类似的，细节请参见第 1 章和第 11 章). 在这幅图中，水平线上的实心方块对应的$\mathcal{W}$的元素是零. 因此我们看到每一行个数不等的非零元素位于不同的位置(与时间段相关). 显然，$\mathcal{W}$的第 0，8，12，14 和 15 个元素为

$$\mathcal{W}_{0\cdot}^{\mathrm{T}}=\left[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},\underbrace{0,\cdots,0}_{11个0}\right],\quad \mathcal{W}_{8\cdot}^{\mathrm{T}}=\left[-\frac{1}{2},-\frac{1}{2},\frac{1}{2},\frac{1}{2},\underbrace{0,\cdots,0}_{12个0}\right],$$

$$\mathcal{W}_{12\cdot}^{T} = \Big[\underbrace{-\frac{1}{\sqrt{8}},\cdots,-\frac{1}{\sqrt{8}}}_{4个},\underbrace{\frac{1}{\sqrt{8}},\cdots,\frac{1}{\sqrt{8}}}_{4个},\underbrace{0,\cdots,0}_{8个0}\Big],$$

$$\mathcal{W}_{14\cdot}^{T} = \Big[\underbrace{-\frac{1}{4},\cdots,-\frac{1}{4}}_{8个},\underbrace{\frac{1}{4},\cdots,\frac{1}{4}}_{8个}\Big],\quad \mathcal{W}_{15\cdot}^{T} = \Big[\underbrace{\frac{1}{4},\cdots,\frac{1}{4}}_{16个}\Big].$$

剩下的11行是将上面这几行平移后得到的结果：

$$\mathcal{W}_{1\cdot} = \mathcal{T}^{2}\mathcal{W}_{0\cdot}\quad \mathcal{W}_{2\cdot} = \mathcal{T}^{4}\mathcal{W}_{0\cdot}\quad \cdots\quad \mathcal{W}_{7\cdot} = \mathcal{T}^{14}\mathcal{W}_{0\cdot}$$
$$\mathcal{W}_{9\cdot} = \mathcal{T}^{4}\mathcal{W}_{8\cdot}\quad \mathcal{W}_{10\cdot} = \mathcal{T}^{8}\mathcal{W}_{8\cdot}\quad \mathcal{W}_{11\cdot} = \mathcal{T}^{12}\mathcal{W}_{8\cdot}$$
$$\mathcal{W}_{13\cdot} = \mathcal{T}^{8}\mathcal{W}_{12\cdot}$$

通过构造，$\mathcal{W}$的每一行都具有单位能量，并且由图容易看到任意不同的两行其内积均为零，因此$\mathcal{W}$就是一个规范正交阵.

练习[58] 对于$N=2^J$，给出哈尔离散小波变换的矩阵$\mathcal{W}$元素的形式化定义，并简要说明为什么$\mathcal{W}$是一个规范正交矩阵. ◁

现在让我们来定义尺度这个概念的确切含义. 对于一个正整数λ，令

$$\overline{X}_t(\lambda) \equiv \frac{1}{\lambda}\sum_{l=0}^{\lambda-1} X_{t-l}$$

表示λ个从下标$t-\lambda+1$到t这几个相邻数据的平均(注意到$\overline{X}_t(1)=X_t$，可认为是单点平均，$\overline{X}_{N-1}(N)=\overline{X}$，即所有数据样本平均). 我们称$\overline{X}_t(\lambda)$是在尺度$\lambda$上从时间$t-\lambda+1$到$t$的样本平均. 由于$\mathbf{W}=\mathcal{W}\mathbf{X}$，考虑$\mathcal{W}$的每一行，我们可以写成

$$\mathbf{W} = \begin{bmatrix} W_0 \\ \vdots \\ W_7 \\ W_8 \\ \vdots \\ W_{11} \\ W_{12} \\ W_{13} \\ W_{14} \\ W_{15} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}(X_1 - X_0) \\ \vdots \\ \frac{1}{\sqrt{2}}(X_{15} - X_{14}) \\ \frac{1}{2}(X_3 + X_2 - X_1 - X_0) \\ \vdots \\ \frac{1}{2}(X_{15} + X_{14} - X_{13} - X_{12}) \\ \frac{1}{\sqrt{8}}(X_7 + \cdots + X_4 - X_3 - \cdots - X_0) \\ \frac{1}{\sqrt{8}}(X_{15} + \cdots + X_{12} - X_{11} - \cdots - X_8) \\ \frac{1}{4}(X_{15} + \cdots + X_8 - X_7 - \cdots - X_0) \\ \frac{1}{4}(X_{15} + \cdots + X_0) \end{bmatrix}.$$

使用$\overline{X}_t(\lambda)$的定义，可以重新将W_n写为

$$W_0 = \frac{1}{\sqrt{2}}[\overline{X}_1(1) - \overline{X}_0(1)],\cdots,W_7 = \frac{1}{\sqrt{2}}[\overline{X}_{15}(1) - \overline{X}_{14}(1)],$$
$$W_8 = \overline{X}_3(2) - \overline{X}_1(2),\cdots,W_{11} = \overline{X}_{15}(2) - \overline{X}_{13}(2),$$
$$W_{12} = \sqrt{2}[\overline{X}_7(4) - \overline{X}_3(4)],W_{13} = \sqrt{2}[\overline{X}_{15}(4) - \overline{X}_{11}(4)],$$
$$W_{14} = 2[\overline{X}_{15}(8) - \overline{X}_7(8)],$$
$$W_{15} = 4\overline{X}_{15}(16).$$

注意前 8 个离散小波变换系数 W_0，…，W_7 与$\{X_t\}$在尺度 1 上的相邻平均的差分成正比；接下来的 4 个 W_8，…，W_{11} 与尺度 2 上的相邻平均的差分成正比；W_{12}，W_{13} 与尺度 4 上的相邻平均的差分成正比；W_{14} 与尺度 8 上的相邻平均的差分成正比；最后一个系数 W_{15} 与全部数据的平均成正比.

对于一般的 $N=2^J$ 和哈尔离散小波变换以及本章所讲的其他离散小波变换，$\mathcal{W}$的元素可以按特定方式排列，使得离散小波变换的前 $N/2$ 个系数与尺度 1 的相邻平均差分有关，接下来的 $N/4$ 个系数与尺度 2 的相邻平均差分相关，依次类推，直到系数 W_{N-4} 和 W_{N-3} 与尺度 $N/4$ 的相邻平均差分相关，W_{N-2} 与尺度 $N/2$ 的相邻平均差分相关，最后一个系数 W_{N-1} 与全部数据的平均成正比. 在尺度 τ_j 上，$j=1$，…，J，共有 $N/(2\tau_j)$个离散小波变换系数与差分相关(注意 $\tau_1=1$ 和 $\tau_J=N/2$). 此外，最后一个系数 W_{N-1} 与尺度 N 的平均相关. 前 $N-1$ 个与各尺度的相邻平均差分相关的系数称为*小波系数*，而系数 W_{N-1} 则称为*尺度系数*. 关于某一尺度的$\mathcal{W}$的各行是其中任一行的循环平移. 对于尺度 τ_j，相邻行之间平移的跨度为 $2\tau_j=2^j$.

重要的是注意到，τ_j 是没有单位的标准化尺度. 在实际应用中，我们必须考虑采样间隔Δt，这时标准尺度 τ_j 就相当于实际的物理尺度 $\tau_j\Delta t$. 举个例子，在 4.10 节我们将会讨论心电图记录的时间序列，其采样率为每秒 180 个样本，因此采样间隔是 $\Delta t=1/180$ 秒. 物理尺度 $\tau_j\Delta t$ 因此以秒为单位测量，例如标准尺度的 $\tau_3=4$ 对应的物理尺度为 $\tau_3\Delta t=1/45$ 秒.

每个尺度上的每个小波系数在时间上也具有局部化特性. 在上面的例子中，W_0 的值只涉及 $t=0$ 和 1 处的时间序列值，而 W_8 的值和 $t=0$，1，2，3，的时间序列值都有关. 作为对比，回想规范正交离散傅里叶变换的变换系数在任何意义下都不是局部化的，这是以规范正交离散傅里叶变换为代表的全局变换和局部化离散小波变换的重大区别. 小波系数的概念只和$\{X_t\}$在不同部分的(加权)平均的(各阶)差分相关，$\{X_t\}$按时间段集中不仅对哈尔离散小波变换是特别的，而且也是所有小波变换的基础(在第 1 章我们已经论证了这一点).

作为不是哈尔离散小波变换的离散小波变换的例子，图 60 表示相应 D(4)小波的 $N=16$ 点离散小波变换的矩阵$\mathcal{W}$，D(4)是 4 项成员离散 Daubechies 小波(这类小波将在 4.8 节中正式定义；哈尔小波是 2 项成员小波). 和前面一样，矩阵$\mathcal{W}$的前 8 行对应于尺度为 1 的情形(参见这一节评论与扩展的第 2 条). 这些行的每个元素基于如下四个非零值：

$$h_0=\frac{1-\sqrt{3}}{4\sqrt{2}},\quad h_1=\frac{-3+\sqrt{3}}{4\sqrt{2}},\quad h_2=\frac{3+\sqrt{3}}{4\sqrt{2}},\quad h_3=\frac{-1-\sqrt{3}}{4\sqrt{2}}. \tag{59a}$$

例如，我们有

$$\mathcal{W}_{0\bullet}^{T}=[h_1,h_0,\underbrace{0,\cdots,0}_{12\text{个}0},h_3,h_2] \quad 和 \quad \mathcal{W}_{1\bullet}^{T}=[h_3,h_2,h_1,h_0,\underbrace{0,\cdots,0}_{12\text{个}0}];$$

即$\mathcal{W}_{1\bullet}=\mathcal{T}^2\mathcal{W}_{0\bullet}$. $\mathcal{W}$的规范正交性要求

$$\|\mathcal{W}_{0\bullet}\|^2=h_0^2+h_1^2+h_2^2+h_3^2=1 \quad 和 \quad \langle\mathcal{W}_{0\bullet},\mathcal{W}_{1\bullet}\rangle=h_0h_2+h_1h_3=0, \tag{59b}$$

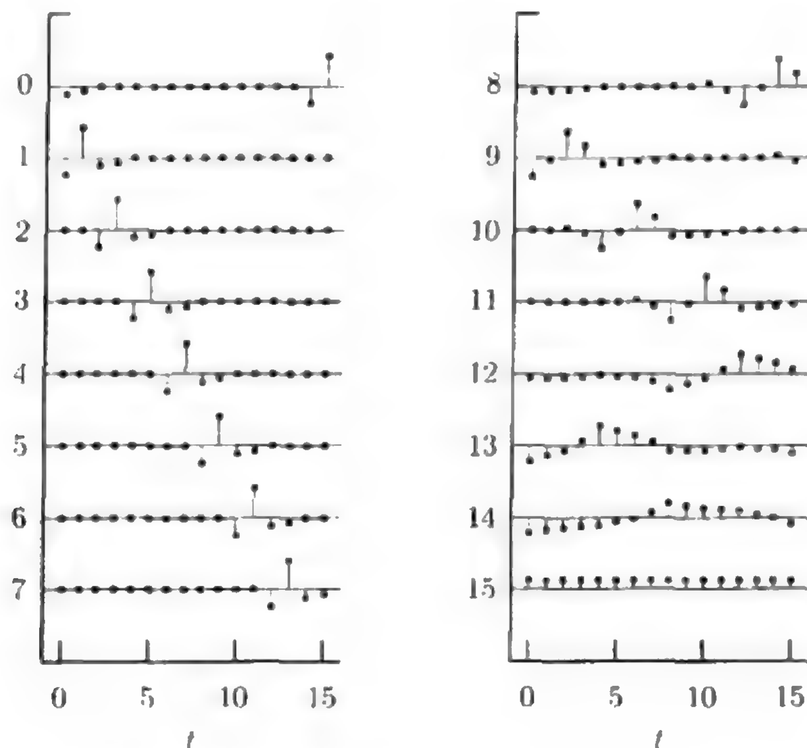


图60　基于D(4)小波的离散小波变换的矩阵的行向量$\mathcal{W}_{n\cdot}^{T}$，其中$N=16$，$n=0$到7(左图从上到下)，$n=8$到15(右图)

这一点读者容易验证.

现在，在单位尺度上的哈尔离散小波变换只基于两个值，即$\frac{1}{\sqrt{2}}$和$-\frac{1}{\sqrt{2}}$. 正如我们前面所指出的，单位尺度上的每个哈尔小波系数可以通过相邻"单点平均"之间做一阶向后差分乘以一个适当尺度得到，使得$W_n \propto X_{2n+1}-X_{2n}$对$n=0, \cdots, 7$成立. 相应地，在单位尺度上的D(4)小波系数也可以通过做相邻两点加权平均的二阶向后差分得到，即先通过式$Y_t \equiv aX_t+bX_{t-1}$得到加权平均，然后再对Y_t做二阶向后差分，也就是做一阶向后差分后再做一阶向后差分. 于是，如果$Y_t^{(1)}=Y_t-Y_{t-1}$表示$\{Y_t\}$的一阶向后差分，则二阶向后差分可表示为

$$Y_t^{(2)} \equiv Y_t^{(1)}-Y_{t-1}^{(1)}=Y_t-2Y_{t-1}+Y_{t-2}.$$

选择特定的a和b，单位尺度上的D(4)小波系数由下式给出：

$$\begin{aligned} W_n &= Y_{2n+1}-2Y_{2n}+Y_{2n-1} \\ &= aX_{2n+1}+bX_{2n}-2(aX_{2n}+bX_{2n-1})+aX_{2n-1}+bX_{2n-2} \\ &= aX_{2n+1}+(b-2a)X_{2n}+(a-2b)X_{2n-1}+bX_{2n-2} \\ &\equiv h_0X_{2n+1}+h_1X_{2n}+h_2X_{2n-1}+h_3X_{2n-2}. \end{aligned} \tag{60}$$

其中$n=0, \cdots, 7$(这里我们必须定义$X_{-1}\equiv X_{15}$和$X_{-2}\equiv X_{14}$，即假设序列是循环的). 式(59b)的两个条件可以用来求得a和b的解，由此得到在前面式(59a)给

出的 h_0，h_1，h_2 和 h_3，这是解的一种结果(参见练习[4.1])。

现在，让我们直观地比较一下哈尔离散小波变换的 $\mathcal{W}$ 相应单位尺度上(图 57 中 $n=0$ 到 7)的行向量和对应尺度 2 上(图 57 中 $n=8$ 到 11)的行向量. 对应尺度 2 的行向量的非零元素都是单位尺度上的行向量的非零元素的"伸展和压缩"(在更高的尺度上也是如此). 有了这种思想(再稍微发挥一下想象力!)，观察一下图 60，我们会发现 D(4)小波也有同样的模式. 我们将在 4.5 节和 4.6 节中看到，尺度大于等于 2 时对应的行向量的非零元素确实是由单位尺度的行上的非零元素的"伸展和压缩"得到的("伸展和压缩"的相应定义也在这两节中讨论).

我们上面看到的对于哈尔离散小波变换和 D(4) 离散小波变换的模式对其他的 Daubechies 小波也是成立的(详见 4.8 节). 因此，对于偶数 L 和 $L<N$，让 h_0，h_1，…，h_{L-1} 代表 $\mathcal{W}$ 的对应单位尺度的行向量上的非零元素. 单位尺度上每一个小波系数可以通过对相邻 $(L/2)$ 个数据点的加权平均做 $(L/2)$ 阶向后差分再乘一个因子得到. 知道了 $L/2$ 个权重系数，就可以知道所有的 h_l，h_l 可以通过类似式(59b) 由 $L/2$ 个条件确定，即满足 $\sum h_l^2=1$ 且

$$\begin{aligned} h_0h_2+h_1h_3+\cdots+h_{L-3}h_{L-1}&=0\\ h_0h_4+h_1h_5+\cdots+h_{L-5}h_{L-1}&=0\\ &\vdots\\ h_0h_{L-2}+h_1h_{L-1}&=0 \end{aligned} \tag{61a}$$

(即 $\{h_l\}$ 与它自身的偶数平移是正交的). 这 $L/2$ 个方程的解并不是唯一的，因此还需要其他条件(比如极值相位或最接近对称，这将在 4.8 节中讨论)的限制来得到唯一解. 单位尺度上的非零元素可以通过适当的"伸展和压缩"唯一地确定更高尺度上的行向量.

现在我们把向量 $\mathbf{W}$ 分解成 $J+1$ 个子向量. 前 J 个子向量记为 $\mathbf{W}_j$，$j=1$，…，J，第 j 个这样的子向量包含了全部关于尺度 τ_j 的离散小波变换系数. 注意 $\mathbf{W}_j$ 是 $N/2^j$ 个元素的列向量. 最后的子向量 $\mathbf{V}_J$ 只包含尺度系数 W_{N-1}. 之后，能再写为

$$\mathbf{W}=\begin{bmatrix}\mathbf{W}_1\\ \mathbf{W}_2\\ \vdots\\ \mathbf{W}_J\\ \mathbf{V}_J\end{bmatrix}. \tag{61b}$$

当 $N=2^J=16$ 时，$J=4$，我们有

$$\begin{aligned} \mathbf{W}_1^T&=[W_0,W_1,W_2,W_3,W_4,W_5,W_6,W_7]\\ \mathbf{W}_2^T&=[W_8,W_9,W_{10},W_{11}]\\ \mathbf{W}_3^T&=[W_{12},W_{13}]\\ \mathbf{W}_4^T&=[W_{14}]\\ \mathbf{V}_4^T&=[W_{15}]. \end{aligned} \tag{61c}$$

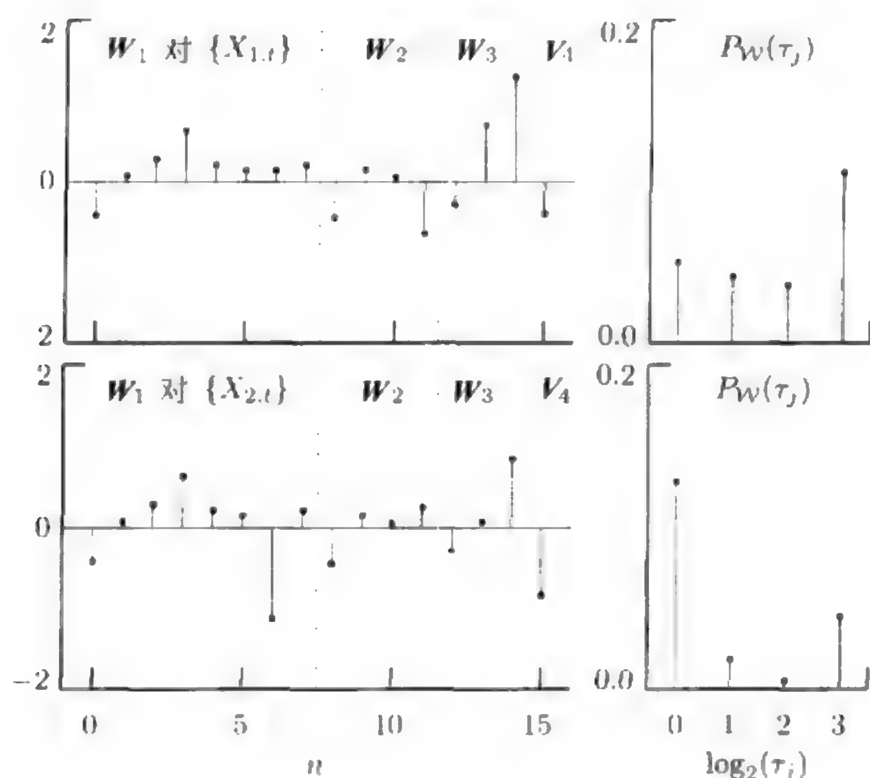


图 62　图 42 中的两个时间序列的哈尔离散小波变换．离散小波变换系数 $\boldsymbol{W}$ 在左图中给出（相应的离散经验功率谱在右图给出）．细的点划线描出了子向量 $\boldsymbol{W}_1$，$\boldsymbol{W}_2$，$\boldsymbol{W}_3$，$\boldsymbol{W}_4$ 和 $\boldsymbol{V}_4$（见式(61c)）；$\boldsymbol{W}_4$ 在 $\boldsymbol{W}_3$ 和 $\boldsymbol{V}_4$ 之间，不过由于空间不够，我们没有标出

我们现在可以将保持能量的条件写成

$$\|\boldsymbol{X}\|^2 = \|\boldsymbol{W}\|^2 = \sum_{j=1}^{J}\|\boldsymbol{W}_j\|^2 + \|\boldsymbol{V}_J\|^2,$$

于是 $\|\boldsymbol{W}_j\|^2$ 表示尺度 τ_j 上的变化对 $\{X_t\}$ 的能量的贡献．练习[97]建立了对使用 Daubechies 小波构成的全部离散小波变换都成立的等式 $W_{N-1}/\sqrt{N}=\overline{X}$．由于这个结果意味着 $\|\boldsymbol{V}_J\|^2=N\overline{X}^2$，因此我们能将样本方差分解为

$$\hat{\sigma}_X^2 = \frac{1}{N}\|\boldsymbol{X}\|^2 - \overline{X}^2 = \frac{1}{N}\|\boldsymbol{W}\|^2 - \overline{X}^2 = \frac{1}{N}\sum_{j=1}^{J}\|\boldsymbol{W}_j\|^2, \tag{62}$$

因此，$\|\boldsymbol{W}_j\|^2/N$ 表示 $\{X_t\}$ 尺度 τ_j 的改变对样本方差的贡献．这个样本方差的分解式可以用来把序列 $\{X_t\}$ 的离散小波变换的经验功率谱 $\{P_W(\tau_j):\ \tau_j=1,\ 2,\ 4,\ \cdots,\ N/2\}$ 定义为

$$P_W(\tau_j) \equiv \frac{1}{N}\|\boldsymbol{W}_j\|^2,\quad P_W(\tau_j)\ \text{满足} \sum_{j=1}^{J}P_W(\tau_j) = \hat{\sigma}_X^2,$$

这个式子和式(48c) 给出的规范正交离散傅里叶变换是等效的．

作为一个例子，图 62 左边的图形表示图 42 中的两个 16 点时间序列$\{X_{1,t}\}$和$\{X_{2,t}\}$的哈尔离散小波变换. 图 62 右边的图形表示相应的离散小波变换的经验功率谱. 注意这两个序列第 13 个值的影响的差别只影响$\mathbf{W}_1$，$\mathbf{W}_2$ 和 $\mathbf{W}_3$ 中的一个系数，以及 $\mathbf{W}_4$ 和 $\mathbf{V}_4$ 中的一个系数. 然而，由于谱是由给定尺度的所有点决定的，因此这两个序列的谱在各个尺度上是不同的.

和离散傅里叶变换的经验功率谱对比，离散小波变换的经验功率谱一般在 $\mathbf{X}$ 的循环平移中不是不变的. 作为一个例子，考虑单位变差时间序列 $\mathbf{X}=[0,0,-2,2,0,0,0,0]^T$，其离散哈尔小波变换的经验功率谱由下式给出：

$$P_W(\tau_j)=\begin{cases}1, & \tau_j=1;\\ 0, & \tau_j=2 \text{ 或 } 4.\end{cases} \tag{63a}$$

而循环平移序列$[0,0,0,-2,2,0,0,0]^T$ 的谱为

$$P_W(\tau_j)=\begin{cases}1/2, & \tau_j=1;\\ 1/4, & \tau_j=2 \text{ 或 } 4.\end{cases} \tag{63b}$$

（练习[4.3]将验证上式.）

现在考虑公式(46b) 指出的 $\mathbf{X}$ 的小波重构：

$$\mathbf{X}=\mathcal{W}^T\mathbf{W}=\sum_{n=0}^{N-1}W_n\,\mathcal{W}_{n\bullet}=\sum_{j=1}^{J}\mathcal{W}_j^T\mathbf{W}_j+\mathcal{V}_J^T\mathbf{V}_J, \tag{63c}$$

其中$\mathcal{W}_j$ 和$\mathcal{V}_J$ 是$\mathcal{W}$的行分块矩阵，与 $\mathbf{W}$ 的分块矩阵 $\mathbf{W}_1$，…，$\mathbf{W}_J$，$\mathbf{V}_J$ 对应. 因此$\frac{N}{2}\times N$ 矩阵$\mathcal{W}_1$ 是由从$\mathcal{W}$的 0 到 $n=\frac{N}{2}-1$ 行构成的；$\frac{N}{4}\times N$ 矩阵$\mathcal{W}_2$ 是由从$\mathcal{W}$的 $n=\frac{N}{2}$到 $n=\frac{3N}{4}-1$ 行构成的；依此类推，最后到 $1\times N$ 矩阵$\mathcal{W}_J$ 和$\mathcal{V}_J$，它们是$\mathcal{W}$的最后两行. 因此我们有

$$\mathcal{W}=\begin{bmatrix}\mathcal{W}_1\\ \mathcal{W}_2\\ \vdots\\ \mathcal{W}_J\\ \mathcal{V}_J\end{bmatrix}, \tag{63d}$$

其中$\mathcal{W}_j$ 是$\frac{N}{2^j}\times N$ 矩阵，$j=1$，…，J，而$\mathcal{V}_J$ 是有 N 个元素的一行向量(实际上其所有元素都等于$\frac{1}{\sqrt{N}}$，参见图 57 和图 60 中第 15 行的图形，并见练习[97]).

在 $N=16$ 点离散小波变换的例子中，$\mathcal{W}_1$ 是 8×16 矩阵，它的行是$\mathcal{W}$的前 8 行；即

$$\mathcal{W}_1=[\mathcal{W}_{0\bullet},\mathcal{W}_{1\bullet},\mathcal{W}_{2\bullet},\mathcal{W}_{3\bullet},\mathcal{W}_{4\bullet},\mathcal{W}_{5\bullet},\mathcal{W}_{6\bullet},\mathcal{W}_{7\bullet}]^T;$$

类似地，$\mathcal{W}_2$ 是如下的 4×16 矩阵：

$$\mathcal{W}_2=[\mathcal{W}_{8\bullet},\mathcal{W}_{9\bullet},\mathcal{W}_{10\bullet},\mathcal{W}_{11\bullet}]^T,$$

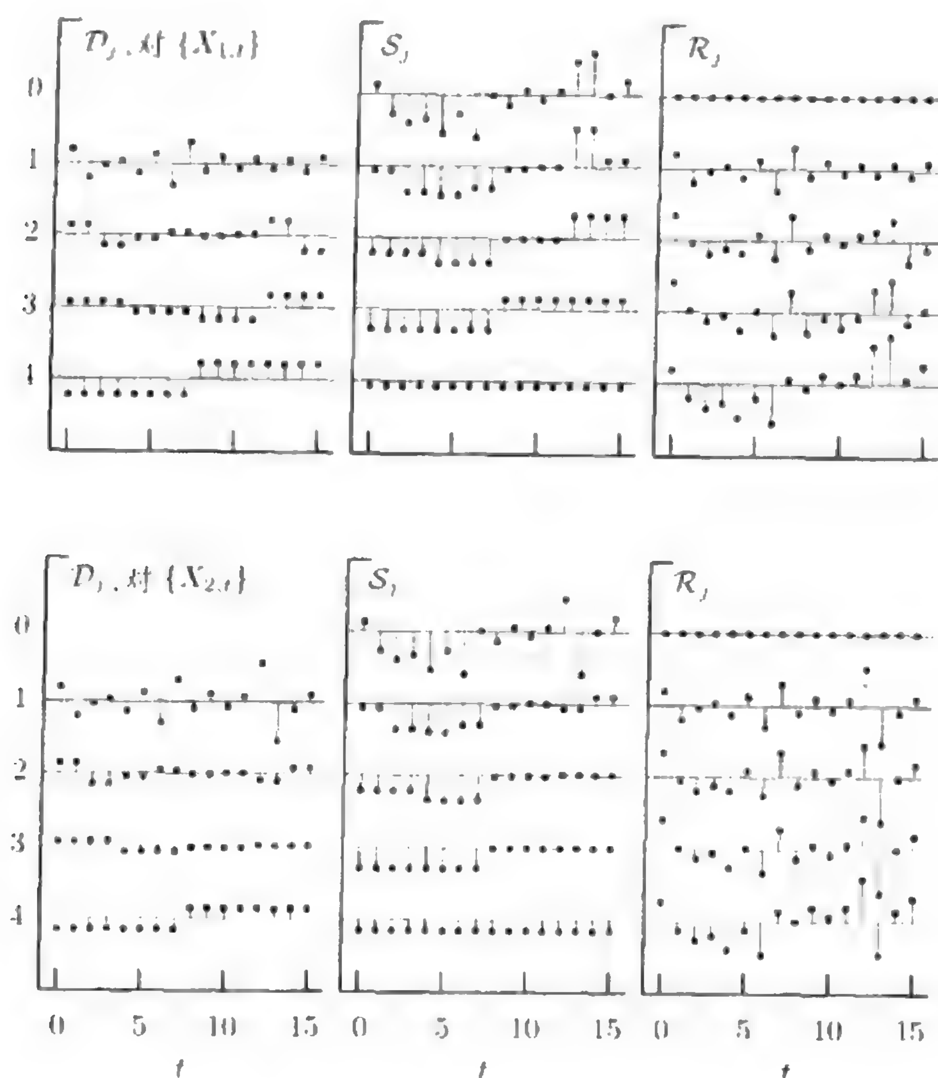


图 64 $\{X_{1,t}\}$(上面的图)和$\{X_{2,t}\}$(下面的图)的 $j=0$ 到 4 的哈尔小波的细节$\mathcal{D}_j$、光滑$\mathcal{S}_j$ 和粗糙$\mathcal{R}_j$. 对任意给定的 j，我们有$\mathcal{S}_j+\mathcal{R}_j=\boldsymbol{X}$. 第 j 层细节可以认为是两个相邻尺度的光滑或者两个相邻尺度的粗糙的差，即$\mathcal{D}_j=\mathcal{S}_{j-1}-\mathcal{S}_j$ 和$\mathcal{D}_j=\mathcal{R}_j-\mathcal{R}_{j-1}$

而$\mathcal{W}_3$，$\mathcal{W}_4$ 和$\mathcal{V}_4$ 是如下给出的 2×16、1×16 和 1×16 矩阵：

$$\mathcal{W}_3=[\mathcal{W}_{12\bullet},\mathcal{W}_{13\bullet}]^{\mathrm{T}},\quad \mathcal{W}_4=\mathcal{W}_{14\bullet}^{\mathrm{T}},\quad \mathcal{V}_4=\mathcal{W}_{15\bullet}^{\mathrm{T}}.$$

现在，我们定义 N 维向量$\mathcal{D}_j\equiv\mathcal{W}_j^{\mathrm{T}}\boldsymbol{W}_j$，$j=1,\cdots,J$，它是 N 维列向量，每个元素与尺度为 τ_j 时 $\boldsymbol{X}$ 中的变化相联系；即 $\boldsymbol{W}_j=\mathcal{W}_j\boldsymbol{X}$ 表示 $\boldsymbol{W}=\mathcal{W}\boldsymbol{X}$ 在尺度 τ_j 上的分量，而$\mathcal{W}_j^{\mathrm{T}}\boldsymbol{W}_j$ 则是 $\boldsymbol{X}=\mathcal{W}^{\mathrm{T}}\boldsymbol{W}$ 在尺度 τ_j 上的分量. 令$\mathcal{S}_J\equiv\mathcal{V}_J^{\mathrm{T}}\boldsymbol{V}_J$，正如式[97]所指出的，$\mathcal{S}_J$ 的每个元素都等于样本平均 $\overline{X}$. 现在我们能写

$$\boldsymbol{X}=\sum_{j=1}^{J}\mathcal{D}_j+\mathcal{S}_J, \tag{64}$$

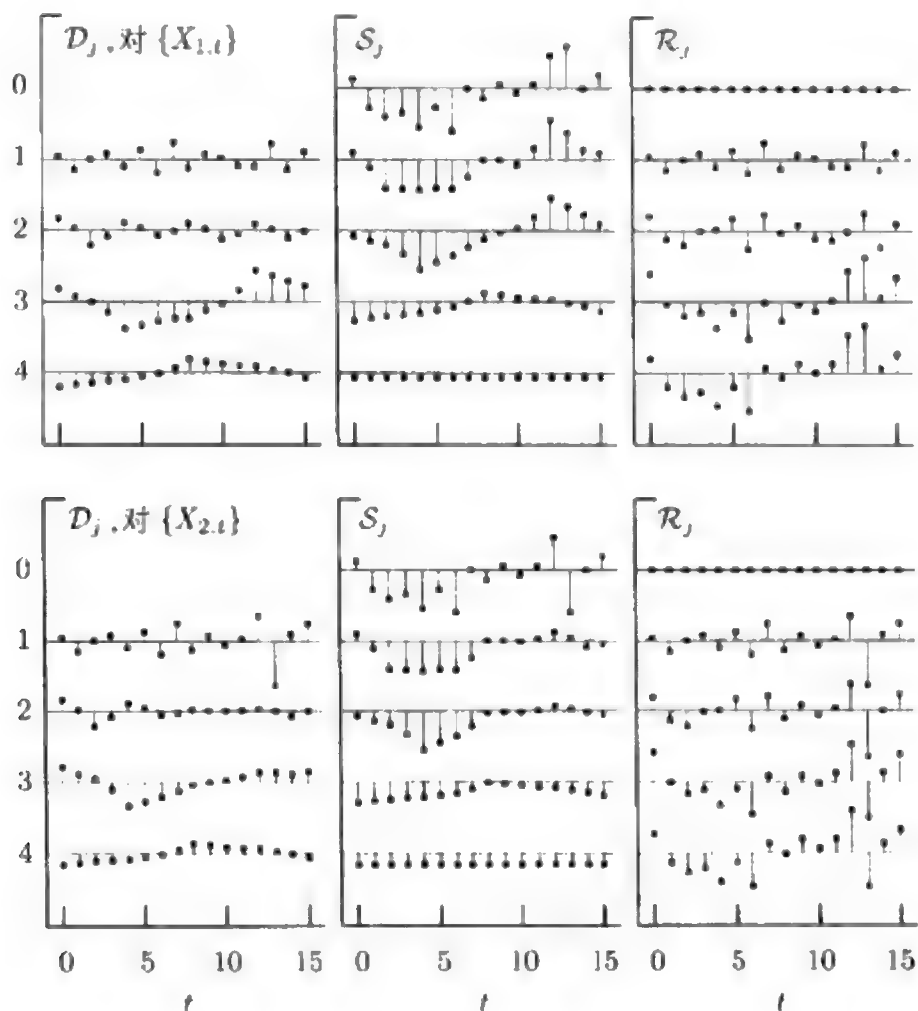


图 65　$\{X_{1,t}\}$(上面的图)和$\{X_{2,t}\}$(下面的图)的 D(4)小波的细节$\mathcal{D}_j$、光滑$\mathcal{S}_j$ 和粗糙$\mathcal{R}_j$. 这幅图和前面哈尔小波的图是相对应的. 比较这两幅图我们可以发现对于一个给定的时间序列来说，它的哈尔和 D(4)的光滑$\mathcal{S}_4$ 完全吻合，因为这两个变换的$\mathcal{V}_4$ 是一样的(由定义知，粗糙$\mathcal{R}_0$ 也是相同的)

它定义了序列 $\boldsymbol{X}$ 的一个多分辨分析(MRA)；即我们将序列 $\boldsymbol{X}$ 表示成了一个常数向量$\mathcal{S}_J$ 和 J 个另外的向量$\mathcal{D}_j$($j=1$，…，J)的和，其中 $\mathcal{D}_j$ 每一个都包含与 $\boldsymbol{X}$ 在某一尺度上的变化有关的一个时间序列. 我们称$\mathcal{D}_j$ 是第 j 层小波细节.

图 64 中顶部与底部两幅图的最左面的图是对图 42 中$\{X_{1,t}\}$和$\{X_{2,t}\}$做哈尔小波变换得到的细节(这里 $J=4$，因为 $N=16=2^4$). 我们来比较一下它们在单位尺度上的细节$\mathcal{D}_1$. 注意这两个序列唯一的不同是第 13 个值，即 $X_{2,13}=-X_{1,13}$，在$\mathcal{D}_1$ 这表现为在 $t=12$ 和 $t=13$ 时两个值不一样. 为看到为什么这种现象发生，注意$\mathcal{D}_1=\mathcal{W}_1^{\mathrm{T}}\boldsymbol{W}_1=\sum_{n=0}^{7}W_n\,\mathcal{W}_{n\cdot}$，并且回忆这两个序列在单位尺度上的小波系数只有 W_6 是不一样的(参见图 62 中最左面的图). 因此细节$\mathcal{D}_j$ 只在 $t=12$ 和$t=13$ 时不一样，因为$\mathcal{W}_{6\cdot}$只在这两个指标(参见图 57)上有非零元素. 图 65 中顶部与底部两幅图的最左面的图表示相应地使用 D(4)小波变换得到的细节图.

这时两个序列的细节$\mathcal{D}_1$在6个指标上都不一样，即在$t=10,\cdots,15$上不同(练习[4.4]说明了为什么会这样)．于是和其他比哈尔小波更宽的小波相比，D(4)的多分辨分析局部性要差些；另一方面，通过图64和图65的比较我们可以发现哈尔小波变换的细节的光滑性比D(4)离散小波变换的细节要差些(注意$\mathcal{D}_4$就是这样)．哈尔小波变换的细节由于受到其块状结构的影响而不能具有很好的光滑性.

$\mathcal{W}$的规范正交性意味着对于$1\leqslant j,k\leqslant J$,

$$\mathcal{D}_j^{T}\mathcal{D}_k=\mathbf{W}_j^{T}\mathcal{W}_j\mathcal{W}_k^{T}\mathbf{W}_k=\begin{cases}\mathbf{W}_j^{T}\mathbf{W}_j, & k=j;\\ 0, & \text{其他}.\end{cases}$$

因此$\|\mathcal{D}_j\|^2=\|\mathbf{W}_j\|^2$，并且，利用式(62)，我们有

$$\hat{\sigma}_X^2=\frac{1}{N}\sum_{j=1}^{J}\|\mathcal{D}_j\|^2,\quad j=1,\cdots,J,$$

其中$\|\mathcal{D}_j\|^2/N$可以解释成N个元素的$\mathcal{D}_j$的样本方差．离散小波变换的经验功率谱$P_W(\tau_j)\equiv\|\mathbf{W}_j\|^2/N$，因此可以用细节来表示，即$P_W(\tau_j)=\|\mathcal{D}_j\|^2/N$.

练习[66] 我们定义了$\mathbf{X}$的样本方差是

$$\hat{\sigma}_X^2=\frac{1}{N}\sum_{t=0}^{N-1}(X_t-\overline{X})^2=\frac{1}{N}\|\mathbf{X}\|^2-\overline{X}^2,$$

其中$\overline{X}$是$\mathbf{X}$元素的样本均值．$\|\mathcal{D}_j\|^2/N$可以认为是$\mathcal{D}_j$的样本方差，但是这和我们定义的$\hat{\sigma}_X^2$是不一致的：我们没有减去$\mathcal{D}_j$的样本均值的平方．请你解释为什么实际上这两个定义是一致的！(可以参见练习[97]的结果.) ◁

回忆我们定义了$\mathcal{S}_J=\mathcal{V}_J^{T}\mathbf{V}_J$，这是一个向量，它的元素是所有的$\overline{X}$，对$0\leqslant j\leqslant J-1$，令

$$\mathcal{S}_j\equiv\sum_{k=j+1}^{J}\mathcal{D}_k+\mathcal{S}_J.\tag{66a}$$

由于对于$j\geqslant1$有

$$\mathbf{X}-\mathcal{S}_j=\sum_{k=1}^{j}\mathcal{D}_k,$$

我们可以证明$\mathcal{S}_j$是光滑化后的$\mathbf{X}$，因为两个向量之差仅涉及尺度$\tau_j=2^{j-1}$的细节并且更小，因此随着指标j的增加，$\mathcal{S}_j$应该看起来更光滑(因为我们注意到在$j=J$时，所有的元素都相等)．我们称$\mathcal{S}_j$是$\mathbf{X}$的第j层小波光滑．类似地，我们定义$\mathbf{X}$的第j层小波粗糙为

$$\mathcal{R}_j\equiv\begin{cases}0, & j=0;\\ \displaystyle\sum_{k=1}^{j}\mathcal{D}_k, & 1\leqslant j\leqslant J,\end{cases}\tag{66b}$$

于是有$\mathbf{X}=\mathcal{S}_j+\mathcal{R}_j$对所有的$j$成立．注意$\mathcal{S}_j-\mathcal{S}_{j+1}=\mathcal{D}_{j+1}$和$\mathcal{R}_{j+1}-\mathcal{R}_j=\mathcal{D}_{j+1}$；即细节是相邻的两个光滑之差或相邻的两个粗糙之差．图64中间和右边一列给出了哈尔小波对序列$\{X_{1,t}\}$(上图)和$\{X_{2,t}\}$(下图)的各个水平的光滑和粗糙；图65则给出了D(4)小波的光滑和粗糙.

4.1 节的关键结论与定义

在限制假设时间序列的长度是 $N=2^J$ 的条件下，$N\times N$ 的规范正交离散小波变换矩阵的行可以分组成 $J+1$ 个子矩阵，相应的离散小波变换系数向量 $\mathbf{W}$ 也被分块：

$$\mathcal{W}\mathbf{X}=\begin{bmatrix}\mathcal{W}_1\\ \mathcal{W}_2\\ \vdots\\ \mathcal{W}_J\\ \mathcal{V}_J\end{bmatrix}\mathbf{X}=\begin{bmatrix}\mathcal{W}_1\mathbf{X}\\ \mathcal{W}_2\mathbf{X}\\ \vdots\\ \mathcal{W}_J\mathbf{X}\\ \mathcal{V}_J\mathbf{X}\end{bmatrix}=\begin{bmatrix}\mathbf{W}_1\\ \mathbf{W}_2\\ \vdots\\ \mathbf{W}_J\\ \mathbf{V}_J\end{bmatrix}=\mathbf{W},$$

其中$\mathcal{W}_j$ 是 $N/2^j\times N$ 维的$\mathcal{V}_J$ 是 $1\times N$ 维的；$\mathbf{W}_j$ 是长为 $N/2^j$ 的列向量；$\mathbf{V}_J$ 包含了 $\mathbf{W}$ 的最后一行元素. 在每个$\mathcal{W}_j$ 中，每一行都是其中任一行的循环平移，并且是两两规范正交的(因为$\mathcal{W}$是规范正交矩阵). 向量 $\mathbf{W}_j$ 的小波系数是在尺度 $\tau_j=2^{j-1}$上对序列相邻的(加权)平均后再做(不同阶的)差分. 而 $\mathbf{V}_J$ 中的尺度系数等于序列 $\mathbf{X}$ 的样本平均 $\overline{X}$ 乘以$\sqrt{N}$. 由于$\mathcal{W}$的规范正交性和 $\mathbf{V}_J$ 的特殊形式，我们可以将 $\mathbf{X}$ 的样本方差(经验功率)分解(分析)成相应尺度 τ_1，…，τ_J 上的分量：

$$\hat{\sigma}_X^2=\frac{1}{N}\|\mathbf{X}\|^2-\overline{X}^2=\frac{1}{N}\|\mathbf{W}\|^2-\overline{X}^2=\frac{1}{N}\sum_{j=1}^{J}\|\mathbf{W}_j\|^2\equiv\sum_{j=1}^{J}P_{\mathcal{W}}(\tau_j),$$

其中序列$\{P_{\mathcal{W}}(\tau_j)\}$是离散小波的经验功率谱. 对$\mathcal{W}$和 $\mathbf{W}$ 也进行相应的矩阵分块，我们可以将 $\mathbf{X}$ 表示成 $J+1$ 个长为 N 的向量的和，其中前 J 个是各尺度上的变化，最后一个向量$\mathcal{S}_J$ 的所有元素都等于样本均值：

$$\mathbf{X}=\mathcal{W}^{\mathrm{T}}\mathbf{W}=[\mathcal{W}_1^{\mathrm{T}},\mathcal{W}_2^{\mathrm{T}},\cdots,\mathcal{W}_J^{\mathrm{T}},\mathcal{V}_J^{\mathrm{T}}]\begin{bmatrix}\mathbf{W}_1\\ \mathbf{W}_2\\ \vdots\\ \mathbf{W}_J\\ \mathbf{V}_J\end{bmatrix}$$

$$=\sum_{j=1}^{J}\mathcal{W}_j^{\mathrm{T}}\mathbf{W}_j+\mathcal{V}_J^{\mathrm{T}}\mathbf{V}_J\equiv\sum_{j=1}^{J}\mathcal{D}_j+\mathcal{S}_J.$$

因为 $\|\mathcal{D}_j\|^2=\|\mathbf{W}_j\|^2$，$j=1$，…，$J$，在各个尺度上的方差分析(ANOVA)可以重新表示为

$$\hat{\sigma}_X^2=\frac{1}{N}\sum_{j=1}^{J}\|\mathcal{D}_j\|^2.$$

细节向量可以逐次相加得到第 j 层光滑和粗糙：

$$\mathcal{S}_j=\sum_{k=j+1}^{J}\mathcal{D}_k+\mathcal{S}_J,\quad \mathcal{R}_j=\sum_{k=1}^{j}\mathcal{D}_k,$$

借助于此，我们有 $\mathbf{X}=\mathcal{S}_j+\mathcal{R}_j$，$j=1$，…，$J$. 第 j 层光滑$\mathcal{S}_j$ 和大于等于 τ_{j+1}的尺度有关，而第 j 层粗糙$\mathcal{R}_j$ 和小于等于 τ_j 的尺度有关.

4.1 节的评论与扩展

[1]对比一下图50和图64我们可以发现，正交离散傅里叶变换和离散小波变换的粗糙与光滑随指标变换呈相反的方向，即规范正交离散傅里叶变换随着 j 的变小而变得更光滑，而离散小波变换则是随着 j 的变小而变得更粗糙.

[2]尺度 τ_j 记号对哈尔小波来说定义是自然的，因为每个哈尔小波系数(除了最后一个)与做 $\tau_j=2^{j-1}$ 点的相邻平均后的差分成正比. 如何对D(4)和其他小波定义尺度我们还不是很清楚，不过通过序列的“等价长度”的概念可以给出尺度的定义——具体细节见4.6节评论与扩展第3条. “等价长度”的概念使得对所有的Daubechies小波都可以定义尺度 τ_j，因此我们可以放心地使用尺度的概念.

[3]注意D(4)小波系数的两步构造，就是将一个滤波器用两个滤波器级联实现，其中一个是2系数的 $\{a_{1,0}=a,\ a_{1,1}=b\}$，这个滤波器实现加权平均过程，另一个滤波器是 $\{a_{2,0}=1,\ a_{2,1}=-2,\ a_{2,2}=1\}$，其功能是实现二阶的向后差分(如何确定 a，b 参见练习[4.1]). 如果用 $\mathbf{X}$ 和 $\mathbf{Y}$ 作为滤波器级联后的输入和输出，那么滤波过程就可用下图表示：

$$\mathbf{X}\longrightarrow\boxed{\{a,b\}}\longrightarrow\boxed{\{1,-2,1\}}\longrightarrow\mathbf{Y}$$

或者等价于单个滤波器的滤波过程如下图：

$$\mathbf{X}\longrightarrow\boxed{\{h_0,h_1,h_2,h_3\}}\longrightarrow\mathbf{Y}$$

其中 h_0，…，h_3 由式(59a)给出.

4.2 小波滤波器

在这一节我们开始精确地定义离散小波变换(到现在为止我们只是通过练习[58]做了这件事，也只定义了哈尔离散小波变换). 我们的定义将形成一种允许 $\mathcal{W}$ 分解成非常稀疏的矩阵的乘积算法. 这个算法就是塔式算法(pyramid algorithm)，在小波文献(1989b)中由Mallat首先提出. 这个算法使得 $\mathbf{W}=\mathcal{W}\mathbf{X}$ 的计算只需要 $O(N)$ 次乘法，而直接计算 $N\times N$ 的矩阵 $\mathcal{W}$ 和向量 $\mathbf{X}$ 的乘积需要 N^2 次乘法(记号 $O(a_N)$ 是指存在一个常数 C 使得对所有的 N，实际的乘法数量小于或等于 Ca_N). 与此相关的思想对于规范正交离散傅里叶变换或快速傅里叶变换也存在，这里使用快速傅里叶变换(FFT)算法需要的乘法次数大约为 $O(N\log_2 N)$，因此离散小波变换的塔式算法比快速傅里叶变换的计算还要快.

我们将用两种方式描述塔式算法，一种是从线性滤波运算的角度，另一种是从矩阵运算的角度. 我们从滤波的角度开始描述，它是关于实值小波滤波器 $\{h_l:\ l=0,\cdots,L-1\}$ 建立的，其中 L 是滤波器的长度且必为偶整数(见练习[69]). 对于具有长度 L 的 $\{h_l\}$，一定有 $h_0\neq 0$ 且 $h_{L-1}\neq 0$. 对 $l<0$ 和 $l\geq L$ 定义 $h_l=0$，于是 $\{h_l\}$ 可

看成是至多有 L 个非零元素的无限序列．一个小波滤波器必须满足下面三个基本性质：

$$\sum_{l=0}^{L-1} h_l = 0; \tag{69a}$$

$$\sum_{l=0}^{L-1} h_l^2 = 1 \tag{69b}$$

和

$$\sum_{l=0}^{L-1} h_l h_{l+2n} = \sum_{l=-\infty}^{\infty} h_l h_{l+2n} = 0 \tag{69c}$$

对所有的非零整数 n 都成立．换句话说，一个小波滤波器其元素和必须是零；必须有单位能量；并且必须与自身的偶数平移正交．第一个条件只是保证小波的基本概念(参见第 1 章)，后面两个条件也是我们在前边讨论的离散小波变换的定性的描述(参见式(61a)及其讨论)．我们称式(69b)和(69c)两个性质一块是小波滤波器的正交性质．事实上在前一节我们已经看到了两个小波滤波器，即哈尔小波滤波器$\{h_0=1/\sqrt{2},\ h_1=-1/\sqrt{2}\}$和式(59a)给出的 $L=4$ 的 D(4)小波滤波器．

练习[69] 假定$\{h_l'\colon l=0,\cdots,L-1\}$是长 L 为奇数的滤波器(这就假定了$h_0'\neq 0$ 且$h_{L-1}'\neq 0$)．对于 $l<0$ 和 $l\geq L$，定义 $h_l'=0$．解释为什么这个滤波器不能满足式(69c)． ◁

令 $H(\cdot)$是$\{h_l\}$的传递函数，即

$$H(f)\equiv\sum_{l=-\infty}^{\infty} h_l \mathrm{e}^{-\mathrm{i}2\pi fl}=\sum_{l=0}^{L-1} h_l \mathrm{e}^{-\mathrm{i}2\pi fl},$$

并且令$\mathcal{H}(\cdot)$表示相应的平方增益函数，即

$$\mathcal{H}(f)\equiv|H(f)|^2.$$

导出与式(69b)和(69c) 等价且形式上更简单的条件是很有用的，而这是借助于平方增益函数$\mathcal{H}(\cdot)$来表示的，即

$$\mathcal{H}(f)+\mathcal{H}\left(f+\frac{1}{2}\right)=2 \tag{69d}$$

对所有的 f 成立．为了建立等价性，首先假设$\{h_l\}$是任一实值滤波器，它的平方增益函数满足上述条件．令

$$h\star h_j\equiv\sum_{l=-\infty}^{\infty} h_l h_{l+j},\quad j=\cdots,-1,0,1,\cdots$$

是这个滤波器的自相关序列(参见式(25b)，回忆由于 h_l 是实数，我们有$h_l^*=h_l$)．式(25c)告诉我们$\{h\star h_j\}$的傅里叶变换是用$|H(f)|^2=\mathcal{H}(f)$给

出的．练习[23b]指出$\{h\star h_{2n}\}$（即只含$\{h\star h_j\}$的偶指数的无限序列）的离散傅里叶变换是用$\frac{1}{2}\left[\mathcal{H}\left(\frac{f}{2}\right)+\mathcal{H}\left(\frac{f}{2}+\frac{1}{2}\right)\right]$给出．对其进行离散傅里叶逆变换（式(35a)）可得

$$h\star h_{2n}=\frac{1}{2}\int_{-1/2}^{1/2}\left[\mathcal{H}\left(\frac{f}{2}\right)+\mathcal{H}\left(\frac{f}{2}+\frac{1}{2}\right)\right]\mathrm{e}^{\mathrm{i}2\pi fn}\mathrm{d}f. \tag{70a}$$

因为我们假定了式(69d)对所有的f成立，当然对所有的$f/2$也成立，因此上式导出了

$$\sum_{l=-\infty}^{\infty}h_l h_{l+2n}=h\star h_{2n}=\int_{-1/2}^{1/2}\mathrm{e}^{\mathrm{i}2\pi fn}\mathrm{d}f=\begin{cases}1, & n=0;\\ 0, & n=\cdots,-2,-1,1,2,\cdots.\end{cases}$$

这就证明了式(69b)和式(69c)必定成立．

练习[70] 为了证明上面的等价条件，现在假设滤波器$\{h_l\}$满足式(69b)和(69c)，证明式(69d)必定是真的． ◁

为了得到单位尺度上的小波系数，我们对时间序列$\{X_t:\ t=0,\ \cdots,\ N-1\}$用$\{h_l\}$进行周期滤波，滤波后的系数每相邻两个只保留一个，这里要求$N\equiv 2^J$，J是某个整数．将用$\{h_l\}$对$\{X_t\}$周期滤波后得到系数记为

$$2^{1/2}\widetilde{W}_{1,t}\equiv\sum_{l=0}^{L-1}h_l X_{t-l\bmod N},\quad t=0,\cdots,N-1. \tag{70b}$$

我们定义单位尺度上的小波系数为

$$W_{1,t}\equiv 2^{1/2}\widetilde{W}_{1,2t+1}=\sum_{l=0}^{L-1}h_l X_{2t+1-l\bmod N},\quad t=0,\cdots,\frac{N}{2}-1. \tag{70c}$$

$N/2$个小波系数的$W_{1,t}$（和$\widetilde{W}_{1,t}$）的两个下标的第一个记录尺度为$\tau_j=2^{j-1}$，因此这里$j=1$对应的是单位尺度．注意到小波系数$\{W_{1,t}\}$是滤波后输出的$\{2^{1/2}\widetilde{W}_{1,t}\}$的$N/2$个奇数指标的值．每相邻两个输出取一个输出的过程称为采样因子为2的下采样或子采样（参见这一节的评论与扩展条目[1]）．式(70c)中那个因子$2^{1/2}$是必需的，目的是为了保持能量不变．

我们现在可以将$\{W_{1,t}\}$的定义和矩阵公式$\mathbf{W}=\mathcal{W}\mathbf{X}$联系起来，其中$\mathbf{W}$（像前面）是对应离散小波变换系数的$N$维列向量，$\mathcal{W}$是定义离散小波变换的$N\times N$的矩阵，$\mathbf{X}$是包含时间序列$\{X_t\}$的$N$维列向量．$\mathbf{W}$的前$N/2$个元素即其子向量$\mathbf{W}_1$，是由元素$W_{1,t}$，$t=0$，$\cdots$，$\frac{N}{2}-1$定义的．由于$\mathbf{W}_1=\mathcal{W}_1\mathbf{X}$，其中$\mathcal{W}_1$是$\frac{N}{2}\times N$阶矩阵，包含$\mathcal{W}$的前$N/2$行，我们如此定义的$\mathbf{W}_1$已经暗含了$\mathcal{W}_1$的定义．为了更清楚地看出这些行的维数，我们将式(70c)重写为

$$W_{1,t}=\sum_{l=0}^{N-1}h_l^{\circ}X_{2t+1-l\bmod N},\quad t=0,\cdots,\frac{N}{2}-1,$$

其中$\{h_l^\circ: l=0, \cdots, N-1\}$是$\{h_l\}$以 N 为周期扩展得到(参见 2.6 节). 对于$0\leqslant t\leqslant \frac{N}{2}-1$,$\mathcal{W}$或$\mathcal{W}_1$ 的第 t 行$\mathcal{W}_{t\cdot}^{\mathrm{T}}$导出

$$W_{1,t} = \mathcal{W}_{t\cdot}^{\mathrm{T}} \boldsymbol{X} = \sum_{l=0}^{N-1} h_l^\circ X_{2t+1-l \bmod N} = \sum_{l=0}^{N-1} h_{2t+1-l \bmod N}^\circ X_l. \tag{71a}$$

令 $t=0$, 我们得到

$$W_{1,0} = \mathcal{W}_{0\cdot}^{\mathrm{T}} \boldsymbol{X} = \sum_{l=0}^{N-1} h_{1-l \bmod N}^\circ X_l,$$

因此我们必有

$$\mathcal{W}_{0\cdot}^{\mathrm{T}} = [h_1^\circ, h_0^\circ, h_{N-1}^\circ, h_{N-2}^\circ, \cdots, h_2^\circ]. \tag{71b}$$

从式(71a)可以看出$\mathcal{W}_1$ 的其余的$\frac{N}{2}-1$ 行可以表示为$\mathcal{W}_{0\cdot}^{\mathrm{T}}$的循环平移形式, 即

$$\mathcal{W}_{t\cdot}^{\mathrm{T}} = [\mathcal{T}^{2t} \mathcal{W}_{0\cdot}]^{\mathrm{T}}, \quad t = 1, \cdots, \frac{N}{2}-1,$$

其中$\mathcal{T}$是式(52a) 所描述的 $N\times N$ 循环平移矩阵. 例如, 当 $t=1$ 时, 我们得到

$$\mathcal{W}_{1\cdot}^{\mathrm{T}} = [h_3^\circ, h_2^\circ, h_1^\circ, h_0^\circ, h_{N-1}^\circ, h_{N-2}^\circ, \cdots, h_4^\circ].$$

现在需要证明$\mathcal{W}_1$ 的行由 $N/2$ 个正交向量组成. 当 $L\leqslant N$ 时, 周期化的滤波器取简单形式

$$h_l^\circ = \begin{cases} h_l, & 0 \leqslant l \leqslant L-1; \\ 0, & L \leqslant l \leqslant N-1. \end{cases}$$

因此$\mathcal{W}_1$ 的第一行就是

$$\mathcal{W}_{0\cdot}^{\mathrm{T}} = [h_1, h_0, \underbrace{0, \cdots, 0}_{N-L\text{个}0}, h_{L-1}, \cdots, h_2]. \tag{71c}$$

在这种情况下, 由于$\mathcal{T}^{2t}$是规范正交变换, 我们有

$$\langle \mathcal{W}_{t\cdot}, \mathcal{W}_{t\cdot} \rangle = \| \mathcal{W}_{t\cdot} \|^2 = \| \mathcal{T}^{2t} \mathcal{W}_{0\cdot} \|^2 = \| \mathcal{W}_{0\cdot} \|^2 = \sum_{l=0}^{L-1} h_l^2 = 1,$$

因为一个小波滤波器具有单位能量(式(69b)). 对于 $t'\neq t$, 我们还有

$$\langle \mathcal{W}_{t'\cdot}, \mathcal{W}_{t\cdot} \rangle = \mathcal{W}_{t'\cdot}^{\mathrm{T}} \mathcal{W}_{t\cdot} = \mathcal{W}_{0\cdot}^{\mathrm{T}} \mathcal{T}^{-2t'} \mathcal{T}^{2t} \mathcal{W}_{0\cdot} = \sum_{l=0}^{L-1} h_l h_{l+2(t-t')} = 0,$$

因为一个小波滤波器是和自身的偶数平移正交的(式(69c)). 因此, 当 $L\leqslant N$ 时, $\mathcal{W}_1$ 中行的规范正交性的证明直接由小波滤波器的三条基本性质中的两条直接推

出(求和为零-式(69a)-的性质这里并不需要).

上面的证明不能轻松地推广到 $L>N$ 的情形，但是下面的方法对 $L\leqslant N$ 和 $L>N$ 都是有效的. 因为 $\mathcal{W}_1$ 的后 $\frac{N}{2}-1$ 行是 $\mathcal{W}_{0\cdot}^{\mathrm{T}}$ 通过偶数次的循环平移得到的，且 $\mathcal{W}_{0\cdot}^{\mathrm{T}}$ 是由式(71b)得到，我们只要证明

$$h^{\circ}\star h_l^{\circ}\equiv\sum_{n=0}^{N-1}h_n^{\circ}h_{n+l\bmod N}^{\circ}=\begin{cases}1, & l=0;\\ 0, & l=2,4,\cdots,N-2,\end{cases}\tag{72}$$

就可以证明规范正交性了. 如 2.5 节所讨论的，我们有

$$\{h_l^{\circ}:l=0,\cdots,N-1\}\leftrightarrow\left\{H\left(\frac{k}{N}\right):k=0,\cdots,N-1\right\},$$

因此 $\{h^{\circ}\star h_l^{\circ}\}\leftrightarrow\left\{\left|H\left(\frac{k}{N}\right)\right|^2=\mathcal{H}\left(\frac{k}{N}\right)\right\}$(参见式(37e)). 逆傅里叶变换关系导出

$$\begin{aligned}h^{\circ}\star h_l^{\circ}&=\frac{1}{N}\sum_{k=0}^{N-1}\mathcal{H}\left(\frac{k}{N}\right)e^{i2\pi lk/N}\\&=\frac{1}{N}\left(\sum_{k=0}^{\frac{N}{2}-1}\mathcal{H}\left(\frac{k}{N}\right)e^{i2\pi lk/N}+\sum_{k=0}^{\frac{N}{2}-1}\mathcal{H}\left(\frac{k}{N}+\frac{1}{2}\right)e^{i2\pi l\left(\frac{k}{N}+\frac{1}{2}\right)}\right).\end{aligned}$$

练习[72] 通过对 $l=2t$，$t=0$，…，$\frac{N}{2}-1$，将上式化简成式(72)以完成正交性的证明. ◁

因为我们借助于滤波器的输出定义了单位尺度上的小波系数，现在就能够利用滤波理论同样的工具来解释这些系数. 令

$$\mathcal{X}_k\equiv\sum_{t=0}^{N-1}X_te^{-i2\pi tk/N},\quad k=0,\cdots,N-1$$

是 $\{X_t\}$ 的离散傅里叶变换. 应用帕塞瓦尔(Parseval)定理(式(36h))表示

$$\varepsilon_X=\sum_{t=0}^{N-1}X_t^2=\frac{1}{N}\sum_{k=0}^{N-1}|\mathcal{X}_k|^2,$$

于是 $|\mathcal{X}_k|^2/N$ 定义了在频率 k/N 处的能量谱. 现在考虑用 $\{h_l\}$ 对 $\{X_t\}$ 滤波的结果产生

$$2^{1/2}\widetilde{W}_{1,t}=\sum_{l=0}^{L-1}h_lX_{t-l\bmod N}=\sum_{l=0}^{N-1}h_l^{\circ}X_{t-l\bmod N},\quad t=0,\cdots,N-1,$$

因此序列 $\{2^{1/2}\widetilde{W}_{1,t}\}$ 是通过将 $\{X_t\}$ 和一个长为 N 的滤波器序列 $\{h_l^{\circ}\}$ 进行循环卷积得到的. 这个序列的离散傅里叶变换是 $\left\{H\left(\frac{k}{N}\right)\right\}$. 利用式(37b)再考虑循环卷积的离散傅里叶变换得到

$$\{2^{1/2}\widetilde{W}_{1,t}\}\leftrightarrow\left\{H\left(\frac{k}{N}\right)\mathcal{X}_k\right\},$$

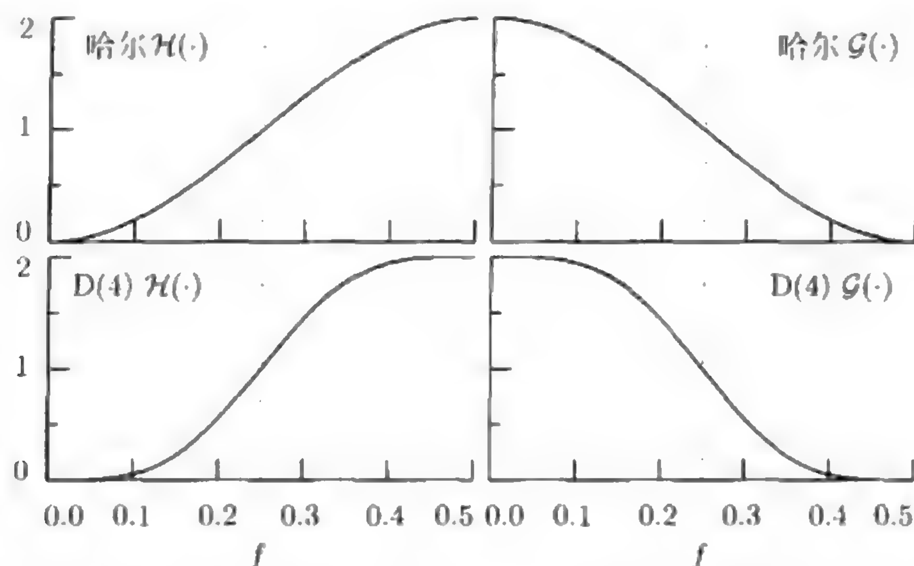


图 73 哈尔小波滤波器的平方增益函数(左上)，哈尔尺度滤波器的平方增益函数(右上)，D(4)小波滤波器的平方增益函数(左下)，D(4)尺度滤波器的平方增益函数(右下)．点划线描绘的是频率 $f=1/4$ 处，即小波滤波器通带的下端，也即尺度滤波器通带的上端

并且由帕塞瓦尔定理转而产生

$$2\sum_{t=0}^{N-1}\widetilde{W}_{1,t}^2=\frac{1}{N}\sum_{k=0}^{N-1}\left|H\left(\frac{k}{N}\right)\mathcal{X}_k\right|^2=\frac{1}{N}\sum_{k=0}^{N-1}\left|H\left(\frac{k}{N}\right)\right|^2|\mathcal{X}_k|^2=\frac{1}{N}\sum_{k=0}^{N-1}\mathcal{H}\left(\frac{k}{N}\right)|\mathcal{X}_k|^2.$$

上式说明从滤波器输出的序列在频率 k/N 处的能量谱由 $\mathcal{H}\left(\frac{k}{N}\right)|\mathcal{X}_k|^2/N$ 给出，且刚好是输入滤波器的序列在频率 k/N 处的能量谱的 $\mathcal{H}\left(\frac{k}{N}\right)$ 倍.

对 $\{X_t\}$ 进行滤波以得到 $\{2^{1/2}\widetilde{W}_{1,t}\}$ 的效果能够作为频率 k/N 研究 $\mathcal{H}\left(\frac{k}{N}\right)$ 的估值. 图 73 左面一列图形分别是对于哈尔和 D(4)小波滤波器的平方增益函数（因为平方增益函数满足 $\mathcal{H}(-f)=\mathcal{H}(f)$ 且以 1 为周期，因此我们只绘了 $0\leqslant f\leqslant\frac{1}{2}$ 上的图形）. 从图上可以看出这些小波滤波器可以认为是近似的高通滤波器，其带通范围是 $\frac{1}{4}\leqslant|f|\leqslant\frac{1}{2}$，其中 D(4)小波滤波器相对来说是个更好的高通滤波器（从工程应用的角度，我们定义带通的下界是平方增益是其最高平方增益的 $\frac{1}{2}$ 处，即 3dB 处，因为 $10\log_{10}(2)\approx 3\text{dB}$；哈尔和 D(4)的临界频率都在 $f=\frac{1}{4}$ 处，因为两个滤波器都满足 $2\,\mathcal{H}\left(\frac{1}{4}\right)=2=\mathcal{H}\left(\frac{1}{2}\right)$）. 这个结果对所有的 Daubechies 小波滤波器都保持(当 L 增大时高通逼近得以改善-参见图 107)，如同 $\{2^{1/2}\widetilde{W}_{1,t}\}$，因此单位尺度上的小波系数 $\{W_{1,t}\}$ 可以认为是频率标称绝对值范围在 $\left[\frac{1}{4},\ \frac{1}{2}\right]$ 内.

4.2 节的关键结论与定义

由定义可知一个长 L 为偶数的实值小波滤波器$\{h_l: l=0, \cdots, L-1\}$满足下面三个条件：

$$\sum_{l=0}^{L-1} h_l = 0, \quad \sum_{l=0}^{L-1} h_l^2 = 1 \quad 和 \quad \sum_{l=-\infty}^{\infty} h_l h_{l+2n} = 0,$$

其中 $h_l=0$，对 $l<0$ 和 $l\geq L$ 均成立，而 n 为任意非零整数．令$\mathcal{H}(\cdot)$是$\{h_l\}$的平方增益函数，对所有 f，条件

$$\mathcal{H}(f) + \mathcal{H}\left(f+\frac{1}{2}\right) = 2$$

与对$\{h_l\}$要求的后两个条件是等价的(即，单位能量性质和与自身的偶平移正交性质)．单位尺度上的 $N/2$ 个小波系数是如下定义的：

$$W_{1,t} \equiv 2^{1/2}\widetilde{W}_{1,2t+1}, \quad t=0,\cdots,\frac{N}{2}-1,$$

而

$$2^{1/2}\widetilde{W}_{1,t} \equiv \sum_{l=0}^{L-1} h_l X_{t-l \bmod N}, \quad t=0,\cdots,N-1;$$

即用$\{h_l\}$对序列$\{X_t\}$进行滤波，得到$\{2^{1/2}\widetilde{W}_{1,t}\}$，然后再对$\{2^{1/2}\widetilde{W}_{1,t}\}$进行下采样得到小波系数．我们可以不通过$\{2^{1/2}\widetilde{W}_{1,t}\}$直接得到小波系数：

$$W_{1,t} = \sum_{l=0}^{L-1} h_l X_{2t+1-l \bmod N} = \sum_{l=0}^{N-1} h_l^{\circ} X_{2t+1-l \bmod N}, \quad t=0,\cdots,\frac{N}{2}-1,$$

其中$\{h_l^{\circ}\}$是对$\{h_l\}$进行周期化所得序列，其周期为 N．这些系数形成了 $\mathbf{W}=\mathcal{W}\mathbf{X}$ 的前 $N/2$ 个系数，即子向量 $\mathbf{W}_1=\mathcal{W}_1\mathbf{X}$ 的元素，其中$\mathcal{W}_1$ 是$\frac{N}{2}\times N$ 阶矩阵，它的行是$\mathcal{W}$的前 $N/2$ 行．$\mathcal{W}_1$ 的第一行为

$$\mathcal{W}_{0\bullet}^{\mathrm{T}} = [h_1^{\circ}, h_0^{\circ}, h_{N-1}^{\circ}, h_{N-2}^{\circ}, \cdots, h_2^{\circ}],$$

剩下的$\frac{N}{2}-1$ 行可以表示成$\mathcal{W}_{0\bullet}^{\mathrm{T}}$的循环平移，即$\mathcal{W}_{t\bullet}^{\mathrm{T}}=[\mathcal{T}^{2t}\mathcal{W}_{0\bullet}]^{\mathrm{T}}$，$t=1, \cdots, \frac{N}{2}-1$．小波滤波器的两个明确的性质(单位能量性质和与自身的偶平移正交性质)意味着$\mathcal{W}_1$ 的各行构成了 $N/2$ 个元素的规范正交向量组．对 Daubechies 小波系中两个特殊的小波滤波器(哈尔和 D(4) 小波)可以验证这些小波滤波器是高通滤波器，其标称通带为 $|f|\in\left[\frac{1}{4}, \frac{1}{2}\right]$．

4.2 节的评论与扩展

[1] 在式(70c)中，我们对滤波输出$\{\widetilde{W}_{1,t}\}$进行下采样，只保留奇数指标的系数. 在工程文献中普遍是保留偶数指标的系数. 如果偶数指标的系数保留下来，我们定义哈尔小波滤波器是$\{h_0, h_1\}$，那么第一个小波系数就是

$$2^{1/2}\widetilde{W}_{1,0} = h_0 X_0 + h_1 X_{N-1},$$

与前面的定义

$$W_{1,0} = 2^{1/2}\widetilde{W}_{1,1} = h_0 X_1 + h_1 X_0$$

不同. 由于哈尔变换的好处就是没有边界效应，因此保留奇数指标会导致更吸引人的哈尔离散小波变换，这样我们就不用将序列$\{X_t\}$看成是一个周期序列了.

4.3 尺度滤波器

在前面几节中，我们用小波滤波器$\{h_l\}$构造离散小波变换矩阵$\mathcal{W}$的前 $N/2$ 行，这些行组成了式(63d)中表示的$\mathcal{W}$的分块阵$\mathcal{W}_1$. 为了通过塔式算法构造$\mathcal{W}$剩下的 $N/2$ 个行向量，现在我们定义第 2 个滤波器以构造$\frac{N}{2}\times N$ 阶矩阵$\mathcal{V}_1$. 我们将证明这个矩阵的各行和$\mathcal{W}$的后 $N/2$ 个行向量张成相同的子空间. 除去对于 $N=2$ 在式(63d)中使 $J=1$ 的情形，$\mathcal{V}_1$ 的各行通常和$\mathcal{W}$的后 $N/2$ 行是不一样的，但是通过一些运算可以从$\mathcal{V}_1$ 得到$\mathcal{W}$的后 $N/2$ 个行向量.

需要的第二个滤波器$\{g_l\}$是与$\{h_l\}$相应的正交镜像滤波器(QMF)：

$$g_l \equiv (-1)^{l+1} h_{L-1-l}. \tag{75a}$$

为了进一步参考，我们注意相逆的关系：

$$h_l = (-1)^l g_{L-1-l}. \tag{75b}$$

滤波器$\{g_l\}$称为*尺度滤波器*. 让我们考虑两个例子. 由于哈尔小波滤波器是由 $h_0=1/\sqrt{2}$，$h_1=-1/\sqrt{2}$给出的，相应的尺度滤波器的非零系数是

$$g_0 = -h_1 = 1/\sqrt{2}, \quad g_1 = h_0 = 1/\sqrt{2}. \tag{75c}$$

对于 D(4) 尺度滤波器，我们有 $g_0=-h_3$，$g_1=h_2$，$g_2=-h_1$，$g_3=h_0$，由式(59a)可得

$$g_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, g_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, g_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, g_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}. \tag{75d}$$

让我们首先注意关于$\{g_l\}$传递函数的基本事实.

练习[76a] 设$\{h_l\}$是一个小波滤波器，令 $H(\cdot)$是它的传递函数. 如(75a)所定义的那样，令$\{g_l\}$是$\{h_l\}$相应的尺度滤波器. 证明$\{g_l\}$的传递函数 $G(\cdot)$为

$$G(f) \equiv \sum_{l=-\infty}^{\infty} g_l e^{-i2\pi fl} = \sum_{l=0}^{L-1} g_l e^{-i2\pi fl} = e^{-i2\pi f(L-1)} H\left(\frac{1}{2}-f\right), \quad (76a)$$

于是有

$$\mathcal{G}(f) = \mathcal{H}\left(\frac{1}{2}-f\right),$$

其中$\mathcal{G}(f) \equiv |G(f)|^2$ 是平方增益函数. ◁

由于$\mathcal{G}(f)=\mathcal{H}\left(\frac{1}{2}-f\right)$，由此得到式(69d) 指出的条件的另一个方法可以写成

$$\mathcal{G}(f)+\mathcal{G}\left(f+\frac{1}{2}\right)=2 \quad 或 \quad \mathcal{G}(f)+\mathcal{H}(f)=2 \quad 对所有 f \quad (76b)$$

(这里我们利用了$\mathcal{G}(\cdot)$和$\mathcal{H}(\cdot)$都是周期为 1 的偶周期函数的事实). 上式中第二个关系对于图 73 中表示的哈尔与 D(4) 平方增益函数是明显的. 这个关系意味着如果小波滤波器看成是高通滤波器，那个尺度滤波器就可以看成是低通滤波器. 图 73 说明对于哈尔尺度函数和 D(4)尺度函数确实是这样的：这些尺度滤波器可以看成是对通带为 $0 \leqslant |f| \leqslant 1/4$ 的理想低通滤波器的近似，D(4)尺度滤波器是更好的近似. 高通滤波器$\{h_l\}$和低通滤波器$\{g_l\}$有时被称为是半带滤波器，因为他们将频带[0，1/2]对半分割了.

下面建立尺度滤波器的一些基本性质.

练习[76b] 证明尺度滤波器必须满足下面的两式之一：

$$\sum_{l=0}^{L-1} g_l = \sqrt{2} \quad 或者 \quad \sum_{l=0}^{L-1} g_l = -\sqrt{2}.$$

进一步证明

$$\sum_{l=0}^{L-1} g_l^2 = 1, \quad \sum_{l=0}^{L-1} g_l g_{l+2n} = \sum_{l=-\infty}^{\infty} g_l g_{l+2n} = 0$$

对所有的非零整数 n 均成立. ◁

从现在开始我们约定 $\sum_l g_l = \sqrt{2}$，这样就可以将尺度滤波器解释为局部加权平均的系数了(而没有负项).

从上面的练习我们看到，小波滤波器和尺度滤波器都满足规范正交性质(即单位能量性质和与自身偶平移规范正交性质). 因为$\{h_l\}$的规范正交性质是$\mathcal{W}_1$ 各行规范正交的充要条件，和用$\{h_l\}$构造矩阵$\mathcal{W}_1$ 相同的方式我们用$\{g_l\}$构造矩阵$\mathcal{V}_1$，那么可以直接得出，$\mathcal{V}_1$ 的各行也是相互规范正交的. 因此我们可以按如下

方式构造$\mathcal{V}_1$. 用$\{g_l\}$对$\{X_t\}$进行循环卷积滤波后的结果记为

$$2^{1/2}\widetilde{V}_{1,t} \equiv \sum_{l=0}^{L-1} g_l X_{t-l \bmod N}, \quad t=0,\cdots,N-1. \tag{77a}$$

定义第一层尺度系数为

$$V_{1,t} \equiv 2^{1/2}\widetilde{V}_{1,2t+1} = \sum_{l=0}^{L-1} g_l X_{2t+1-l \bmod N} = \sum_{l=0}^{N-1} g_l^{\circ} X_{2t+1-l \bmod N}, \tag{77b}$$

$t=0$，…，$\frac{N}{2}-1$，其中$\{g_l^{\circ}\}$是$\{g_l\}$的周期化序列，周期为 N. 令 $\mathbf{V}_1$ 是长为 $N/2$ 的向量，其第 t 个元素是 $V_{1,t}$. 令$\mathcal{V}_1$ 是$\frac{N}{2}\times N$ 阶矩阵，第一行为

$$[g_1^{\circ}, g_0^{\circ}, g_{N-1}^{\circ}, g_{N-2}^{\circ}, \cdots, g_2^{\circ}] \equiv \mathcal{V}_{0\bullet}^{T}.$$

(参见式(71b))并且其余下的$\frac{N}{2}-1$ 个行向量是$[\mathcal{T}^{2t}\mathcal{V}_{0\bullet}]^T$，$t=1$，…，$\frac{N}{2}-1$. 于是有 $\mathbf{V}_1=\mathcal{V}_1\mathbf{X}$，且$\mathcal{V}_1$ 的行向量构成有 $N/2$ 个元素的规范正交向量组.

我们下面的任务是要证明$\mathcal{V}_1$ 和$\mathcal{W}_1$ 一起可以构成 N 维规范正交向量组. 由于$\mathcal{V}_1$ 和$\mathcal{W}_1$ 的第 t 行分别是$[\mathcal{T}^{2t}\mathcal{V}_{0\bullet}]^T$ 和$[\mathcal{T}^{2t}\mathcal{W}_{0\bullet}]^T$，因此需要证明

$$\langle \mathcal{T}^{2t}\mathcal{V}_{0\bullet}, \mathcal{T}^{2t'}\mathcal{W}_{0\bullet}\rangle = 0, \quad 0 \leqslant t \leqslant t' \leqslant \frac{N}{2}-1. \tag{77c}$$

令 $n=t'-t$，对 $n=0$，…，$\frac{N}{2}-1$，我们有

$$\langle \mathcal{T}^{2t}\mathcal{V}_{0\bullet}, \mathcal{T}^{2t'}\mathcal{W}_{0\bullet}\rangle = \mathcal{V}_{0\bullet}^{T}\mathcal{T}^{-2t}\mathcal{T}^{2t'}\mathcal{W}_{0\bullet} = \mathcal{V}_{0\bullet}^{T}\mathcal{T}^{2n}\mathcal{W}_{0\bullet} = \sum_{l=0}^{N-1} g_l^{\circ} h_{l+2n \bmod N}^{\circ}.$$

当 $L\leqslant N$ 时有 $g_l^{\circ}=g_l$ 且 $h_l^{\circ}=h_l$，上式可化简为

$$\langle \mathcal{T}^{2t}\mathcal{V}_{0\bullet}, \mathcal{T}^{2t'}\mathcal{W}_{0\bullet}\rangle = \sum_{l=0}^{L-1} g_l h_{l+2n}.$$

练习[77]　对于特殊情况 $L\leqslant N$，证明

$$\sum_{l=0}^{L-1} g_l h_{l+2n} = \sum_{l=-\infty}^{\infty} g_l h_{l+2n} = 0$$

对所有的整数 n 都成立(使用一个好的时间参数)，进而建立式(77c). ◁

上式说明尺度滤波器的任一个偶数平移和小波滤波器的任一个偶数平移是正交的.

上面关于规范正交性的讨论不能简单地推广到包含 $L>N$ 的情形. 下面的方法对 $L\leqslant N$ 和 $L>N$ 都适用. 我们需要证明

$$\sum_{l=0}^{N-1} g_l^{\circ} h_{l+2n \bmod N}^{\circ} = g^{\circ} \star h_{2n}^{\circ} = 0, \quad n=0,\cdots,\frac{N}{2}-1, \tag{77d}$$

其中$\{g^{\circ}\star h_l^{\circ}\}$是$\{g_l^{\circ}\}$和$\{h_l^{\circ}\}$的循环互相关(参见式(37c)，注意由于$\{g_l^{\circ}\}$是实值的故有 $g_l^{\circ}=[g_l^{\circ}]^*$). 因为$\{g_l^{\circ}\} \leftrightarrow \left\{G\left(\frac{k}{N}\right)\right\}$，且$\{h_l^{\circ}\} \leftrightarrow \left\{H\left(\frac{k}{N}\right)\right\}$，于是我们有 $\{g^{\circ}\star h_l^{\circ}\} \leftrightarrow \left\{G^*\left(\frac{k}{N}\right)H\left(\frac{k}{N}\right)\right\}$(参见式(37d)).

练习[78] 用离散傅里叶逆变换来证明

$$g^{\circ} \star h^{\circ}_{2n} = \frac{1}{N}\sum_{k=0}^{\frac{N}{2}-1}\left[G^{*}\left(\frac{k}{N}\right)H\left(\frac{k}{N}\right)+G^{*}\left(\frac{k}{N}+\frac{1}{2}\right)H\left(\frac{k}{N}+\frac{1}{2}\right)\right]e^{i4\pi nk/N},$$

并且证明对于所有的 k 有

$$G^{*}\left(\frac{k}{N}\right)H\left(\frac{k}{N}\right)+G^{*}\left(\frac{k}{N}+\frac{1}{2}\right)H\left(\frac{k}{N}+\frac{1}{2}\right)=0$$

成立，且直接建立了式(77d). ◁

用矩阵的记号，$\mathcal{V}_1$ 的每一行和 $\mathcal{W}_1$ 的每一行正交能表示为

$$\mathcal{W}_1\mathcal{V}_1^{T} = \mathcal{V}_1\mathcal{W}_1^{T} = 0_{N/2},$$

其中 $0_{N/2}$ 是 $\frac{N}{2}\times\frac{N}{2}$ 阶零矩阵，它的所有元素是零. 因为我们已经知道 $\mathcal{W}_1\,\mathcal{W}_1^{T}=I_{N/2}$ 和 $\mathcal{V}_1\,\mathcal{V}_1^{T}=I_{N/2}$，其中 $I_{N/2}$ 是 $\frac{N}{2}\times\frac{N}{2}$ 阶单位矩阵，所以 $N\times N$ 阶矩阵

$$\mathcal{P}_1 \equiv \begin{bmatrix}\mathcal{W}_1\\ \mathcal{V}_1\end{bmatrix} \tag{78a}$$

是规范正交阵，因为

$$\mathcal{P}_1\,\mathcal{P}_1^{T} = \begin{bmatrix}\mathcal{W}_1\\ \mathcal{V}_1\end{bmatrix}\begin{bmatrix}\mathcal{W}_1^{T} & \mathcal{V}_1^{T}\end{bmatrix} = \begin{bmatrix}\mathcal{W}_1\mathcal{W}_1^{T} & \mathcal{W}_1\mathcal{V}_1^{T}\\ \mathcal{V}_1\mathcal{W}_1^{T} & \mathcal{V}_1\mathcal{V}_1^{T}\end{bmatrix} = I_N. \tag{78b}$$

由于离散小波变换矩阵 $\mathcal{W}$ 也是规范正交的，$\mathcal{W}$ 和 $\mathcal{P}_1$ 的前 $N/2$ 行是相同的，于是必有 $\mathcal{P}_1$ 的后 $N/2$ 个行向量，即 $\mathcal{V}_1$，和 $\mathcal{W}$ 的后 $N/2$ 个行向量张成相同的子空间. 除了 $N=2$ 的情形，$\mathcal{V}_1$ 和 $\mathcal{W}$ 的后 $N/2$ 行是不同的，但是我们可以通过一系列运算从 $\mathcal{V}_1$ 得到 $\mathcal{W}$ 的后 $N/2$ 个行向量. 下一节将详细讨论这个问题.

4.3 节的关键结论与定义

给定实值小波滤波器 $\{h_l\}$，尺度滤波器如下定义：

$$g_l \equiv (-1)^{l+1}h_{L-1-l}.$$

这个滤波器满足如下条件：

$$\sum_{l=0}^{L-1} g_l = \sqrt{2},\quad \sum_{l=0}^{L-1} g_l^2 = 1,\quad \sum_{l=-\infty}^{\infty} g_l g_{l+2n} = 0,\quad \sum_{l=-\infty}^{\infty} g_l h_{l+2n'} = 0$$

对于所有的非零整数 n 和所有的整数 n' 都成立. 令 $G(\cdot)$ 和 $\mathcal{G}(\cdot)$ 分别为 $\{g_l\}$ 的传递函数和平方增益函数，我们有

$$G(f) = e^{-i2\pi f(L-1)}H\left(\frac{1}{2}-f\right),\quad \mathcal{G}(f)+\mathcal{G}\left(f+\frac{1}{2}\right)=2,\quad \mathcal{G}(f)+\mathcal{H}(f)=2$$

对所有的 f 成立. 在实际应用中，$\{g_l\}$ 是一个低通滤波器，其标准通带为 $[-1/4, 1/4]$.

前 $N/2$ 个第一层尺度系数如下定义：

$$V_{1,t} \equiv 2^{1/2}\widetilde{V}_{1,2t+1}, \quad t = 0,\cdots,\frac{N}{2}-1,$$

而

$$2^{1/2}\widetilde{V}_{1,t} \equiv \sum_{l=0}^{L-1} g_l X_{t-l \bmod N}, \quad t = 0,\cdots,N-1.$$

我们也可以直接求得 $V_{1,t}$ 而不必求 $2^{1/2}\widetilde{V}_{1,t}$：

$$V_{1,t} = \sum_{l=0}^{L-1} g_l X_{2t+1-l \bmod N} = \sum_{l=0}^{N-1} g_l^{\circ} X_{2t+1-l \bmod N}, \quad t = 0,\cdots,\frac{N}{2}-1,$$

其中{}是 $\{g_l\}$ 以 N 为周期周期化的结果. 这些系数就是 $\mathbf{V}_1 = \mathcal{V}_1\mathbf{X}$ 的 $N/2$ 个元素，其中 $\mathcal{V}_1$ 是 $\frac{N}{2}\times N$ 阶矩阵，第一行为 $\mathcal{V}_{0\cdot}^{\mathrm{T}} = [g_1^{\circ}, g_0^{\circ}, g_{N-1}^{\circ}, g_{N-2}^{\circ}, \cdots, g_2^{\circ}]$，余下的 $\frac{N}{2}-1$ 行可以通过 $\mathcal{V}_{0\cdot}^{\mathrm{T}}$ 的循环平移得到，即 $[\mathcal{T}^{2t}\mathcal{V}_{0\cdot}]^{\mathrm{T}}$，$t=1, \cdots, \frac{N}{2}-1$. 尺度滤波器的两个性质(单位能量性质和与自身偶平移规范正交性质)意味着 $\mathcal{V}_1$ 的各行构成了一个规范正交向量组. 另外，尺度滤波器和小波滤波器及他们的偶平移是相互正交的，这说明 $\mathcal{V}_1$ 和 $\mathcal{W}_1$ 是相互正交的；当两个矩阵按如下方式叠堆起来时：

$$\mathcal{P}_1 \equiv \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{V}_1 \end{bmatrix},$$

得到的 $N\times N$ 阶矩阵 $\mathcal{P}_1$ 就是规范正交的.

4.3 节的评论与扩展

[1]另一种定义正交镜像滤波器的方式是 $g_l = (-1)^{l-1}h_{1-l}$，相反的关系是 $h_l = (-1)^l g_{1-l}$. 这个定义用得很普遍，例如 Bruce and Gao(1996a)就是这样定义的. 这样定义的好处是可以把它用在无限长的滤波器上，因为定义中不涉及滤波器的长度 L. 我们的定义(75a)假设了滤波器的长度是有限的. 与式(75b)比较表明，式(75b)中的 g_{L-1-l} 用第二个定义就变成了 g_{1-l}，也就是偶平移了 $-L+2$. 作为结果，给定尺度滤波器 $\{g_l: l=0, \cdots, L-1\}$，我们得到小波滤波器的非零系数为 $\{h_l: l=-L+2, \cdots, 0, 1\}$. 对比另一种方法式(75a)给出 $\{h_l, l=0, \cdots, L-1\}$. 我们选择这种定义方式只是因为这种定义强调了尺度滤波器和小波滤波器在下面要讨论的塔式算法中的地位是等同的；而且这样指标不带负数，容易将矩阵和滤波器联系起来. 从数学的角度来看，这两种定义描述的是同一个概念. (注意，对于哈尔小波滤波器来说这两种定义会导致完全一样的尺度滤波器.)

[2]仔细分析证明式(78b)的 $\mathcal{P}_1$ 是规范正交阵的过程，我们会发现只要 N 是偶数，$\mathcal{P}_1$ 就有定义且是规范正交的，也就是说我们不需要假定 $N=2^J$，但在定义离散小波变换的矩阵 $\mathcal{W}$ 时则必须有这个假设. 这个问题将在 4.7 节进一步讨论，在那里我们会看到 $\mathcal{P}_1$ 是一个“部分离散小波变换”矩阵的例子.

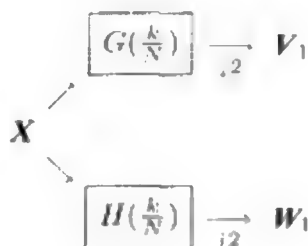


图 80 将 $\boldsymbol{X}$ 分解为 $\boldsymbol{W}_1$ 和 $\boldsymbol{V}_1$ 的流程图. 长为 N 的时间序列用小波滤波器循环滤波，其循环周期为 $N\left(\text{这个滤波器的频域形式为}\left\{H\left(\frac{k}{N}\right): k=0, \cdots, N-1\right\}\right)$. 滤波输出后的奇数指标的元素用来形成第一层小波系数的长为 $N/2$ 的向量 $\boldsymbol{W}_1$. (“$\downarrow 2$”表示因子为 2 的下采样)；类似地，长为 N 的时间序列用尺度滤波器循环滤波，其循环周期为 N，滤波输出后的奇数指标的元素用来形成第一层尺度系数的长为 $N/2$ 的向量 $\boldsymbol{V}_1$

[3]尺度滤波器有时候被称为是父小波滤波器(在法语中是 le père)，而$\{h_l\}$有时被称为是母小波滤波器(在法语中是 la mère). Strichartz(1994)对此发表评论说“……这表现了人们对人类繁殖后代非常无知，小波的产生更像是阿米巴虫的繁殖方式.”因此，我们还是用原来的术语，虽然不够生动，但是可以避免遭到谴责.

4.4 塔式算法的第一步

计算离散小波变换的塔式算法的第一步只是将长为 $N=2^J$ 的时间序列 $\boldsymbol{X}$ 分解成 $N/2$ 个第一层小波系数 $\boldsymbol{W}_1$ 和 $N/2$ 个第一层尺度系数 $\boldsymbol{V}_1$(图 80 从滤波的角度图解了这个变换). 塔式算法还有 $J-1$ 个后继步骤. 对于 $j=2, \cdots, J$，第 j 步变换就是将长为 $N/2^{j-1}$ 的向量 $\boldsymbol{V}_{j-1}$ 分解成向量 $\boldsymbol{W}_j$ 和 $\boldsymbol{V}_j$，其长度都为 $N/2^j$. 在第 j 步分解中，$\boldsymbol{V}_{j-1}$ 就相当于第一步分解中的 $\boldsymbol{X}$：$\boldsymbol{V}_{j-1}$ 的元素分别用$\{h_l\}$和$\{g_l\}$这两个滤波器滤波，其输出分别是 $\boldsymbol{W}_j$ 和 $\boldsymbol{V}_j$. $\boldsymbol{V}_j$ 的元素称为是第 j 层尺度系数，$\boldsymbol{W}_j$ 的元素称为是第 j 层小波系数. 在第 J 步，通过将 $J+1$ 个向量 $\boldsymbol{W}_1, \cdots, \boldsymbol{W}_J$ 和 $\boldsymbol{V}_J$ 合并起来得到离散小波变换系数 $\boldsymbol{W}$(参见式(61b)).

上面描述的塔式算法已经告诉了我们如何计算离散小波变换，但是没有告诉我们离散小波变换从时间序列中可以挖掘到什么信息. 为了理解离散小波变换，我们在这一节和下一节会仔细分析和解释塔式算法的第一步和第二步的一些关键的地方，然后才能在 4.6 节中更好地理解离散小波变换. 这一节我们分析塔式算法第一步的最关键的地方为：由 $\boldsymbol{W}_1$ 和 $\boldsymbol{V}_1$ 重构时间序列 $\boldsymbol{X}$；$\boldsymbol{X}$ 的离散傅里叶变换及 $\boldsymbol{V}_1$ 和 $\boldsymbol{W}_1$ 的离散傅里叶变换之间的关系；阐明尺度系数 $\boldsymbol{V}_1$ 的尺度为 $\lambda_1=2$.

让我们先考虑 $\boldsymbol{X}$ 的重构过程. 因为我们可以用式(78a)中的矩阵 $\mathcal{P}_1$ 来表示塔

式算法的第一步：

$$\mathcal{P}_1\boldsymbol{X}=\begin{bmatrix}\mathcal{W}_1\\\mathcal{V}_1\end{bmatrix}\boldsymbol{X}=\begin{bmatrix}\mathcal{W}_1\boldsymbol{X}\\\mathcal{V}_1\boldsymbol{X}\end{bmatrix}=\begin{bmatrix}\boldsymbol{W}_1\\\boldsymbol{V}_1\end{bmatrix},$$

且因为$\mathcal{P}_1$ 是规范正交矩阵，我们可以通过下式重构 $\boldsymbol{X}$：

$$\boldsymbol{X}=\mathcal{P}_1^{\mathrm{T}}\begin{bmatrix}\boldsymbol{W}_1\\\boldsymbol{V}_1\end{bmatrix}=\begin{bmatrix}\mathcal{W}_1^{\mathrm{T}} & \mathcal{V}_1^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{W}_1\\\boldsymbol{V}_1\end{bmatrix}=\mathcal{W}_1^{\mathrm{T}}\boldsymbol{W}_1+\mathcal{V}_1^{\mathrm{T}}\boldsymbol{V}_1. \tag{81a}$$

回忆 4.1 节中我们定义了第一层细节$\mathcal{D}_1\equiv\mathcal{W}_1^{\mathrm{T}}\boldsymbol{W}_1$，第一层光滑$\mathcal{S}_1$ 使$\mathcal{S}_1+\mathcal{D}_1=\boldsymbol{X}$. 比较上述内容表明我们一定有

$$\mathcal{S}_1=\mathcal{V}_1^{\mathrm{T}}\boldsymbol{V}_1=\mathcal{V}_1^{\mathrm{T}}\mathcal{V}_1\boldsymbol{X},$$

因此 $N\times N$ 阶矩阵$\mathcal{V}_1^{\mathrm{T}}\mathcal{V}_1$ 可以认为是能够从 $\boldsymbol{X}$ 中提取第一层小波光滑的一个算子，类似地，我们有

$$\mathcal{D}_1=\mathcal{W}_1^{\mathrm{T}}\mathcal{W}_1\boldsymbol{X},$$

此式告诉我们如何通过矩阵运算从 $\boldsymbol{X}$ 提取$\mathcal{D}_1$.

更感兴趣的是$\mathcal{D}_1$能通过滤波运算得到. 为了方便起见，我们考虑如下的矩阵$\mathcal{W}_1$，其中$L=4$ 且 $N>4$：

$$\mathcal{W}_1=\begin{bmatrix}h_1 & h_0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & h_3 & h_2\\ h_3 & h_2 & h_1 & h_0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & h_3 & h_2 & h_1 & h_0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & h_3 & h_2 & h_1 & h_0\end{bmatrix}. \tag{81b}$$

由于$\mathcal{D}_1=\mathcal{W}_1^{\mathrm{T}}\boldsymbol{W}_1$，我们有

$$\mathcal{D}_1=\begin{bmatrix}h_1 & h_3 & 0 & \cdots & 0 & 0\\ h_0 & h_2 & 0 & \cdots & 0 & 0\\ 0 & h_1 & h_3 & \cdots & 0 & 0\\ 0 & h_0 & h_2 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & h_1 & h_3\\ 0 & 0 & 0 & \cdots & h_0 & h_2\\ h_3 & 0 & 0 & \cdots & 0 & h_1\\ h_2 & 0 & 0 & \cdots & 0 & h_0\end{bmatrix}\begin{bmatrix}\boldsymbol{W}_{1,0}\\ \boldsymbol{W}_{1,1}\\ \boldsymbol{W}_{1,2}\\ \vdots\\ \boldsymbol{W}_{1,\frac{N}{2}-2}\\ \boldsymbol{W}_{1,\frac{N}{2}-1}\end{bmatrix}. \tag{81c}$$

如果令$\mathcal{D}_{1,t}$是$\mathcal{D}_1$ 的第 t 个元素，则可以写成

$$\mathcal{D}_{1,t}=\begin{cases}h_1\boldsymbol{W}_{1,\frac{t}{2}}+h_3\boldsymbol{W}_{1,\frac{t}{2}+1\bmod\frac{N}{2}}, & t=0,2,\cdots,N-2;\\ h_0\boldsymbol{W}_{1,\frac{t-1}{2}}+h_2\boldsymbol{W}_{1,\frac{t-1}{2}+1\bmod\frac{N}{2}}, & t=1,3,\cdots,N-1.\end{cases}$$

如果比较上式和式(37c)，我们将看到$\mathcal{D}_1$ 的一半元素是通过 $\boldsymbol{W}_1$ 和$\{h_1, h_3\}$的循环互相关形成的，而另一半是通过 $\boldsymbol{W}_1$ 和$\{h_0, h_2\}$的循环互相关形成的. 因此$\mathcal{D}_1$

的元素不是简单的互相关运算而是从 $\mathbf{W}_1$ 的两个互相关运算中交替得到的结果. 由于 $\mathcal{V}_1$ 和 $\mathcal{W}_1$ 的结构类似(g_l 代替了 h_l), 于是第一层光滑 $\mathcal{S}_1$ 可以解释成从两个关于 $\mathbf{V}_1$ 的互相关运算中交替得到的结果. 由于 $\mathbf{X}=\mathcal{D}_1+\mathcal{S}_1$, 我们能使用对于 $\mathcal{D}_1$ 和 $\mathcal{S}_1$ 的这些交插, 对于一般 L 写成

$$X_t = \sum_{l=0}^{\frac{L}{2}-1} h_{2l+1} W_{1,\frac{t}{2}+l \bmod \frac{N}{2}} + \sum_{l=0}^{\frac{L}{2}-1} g_{2l+1} V_{1,\frac{t}{2}+l \bmod \frac{N}{2}},$$

$t=0, 2, \cdots, N-2$, 且

$$X_t = \sum_{l=0}^{\frac{L}{2}-1} h_{2l} W_{1,\frac{t-1}{2}+l \bmod \frac{N}{2}} + \sum_{l=0}^{\frac{L}{2}-1} g_{2l} V_{1,\frac{t-1}{2}+l \bmod \frac{N}{2}},$$

$t=1, 3, \cdots, N-1$.

然而, 第二种关于构造 $\mathcal{D}_1$ 的观点只涉及一个互相关运算构造. 假设我们在 $\mathcal{W}_1^T$ 中再加入 $N/2$ 列向量(每隔一列插入一个新列), 且在 $\mathbf{W}_1$ 中再插入零, 如下式:

$$\mathcal{D}_1 = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & h_0 & h_1 & h_2 & h_3 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & h_0 & h_1 & h_2 & h_3 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & h_0 & h_1 & h_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & h_1 & h_2 & h_3 \\ h_3 & 0 & 0 & 0 & 0 & 0 & \cdots & h_0 & h_1 & h_2 \\ h_2 & h_3 & 0 & 0 & 0 & 0 & \cdots & 0 & h_0 & h_1 \\ h_1 & h_2 & h_3 & 0 & 0 & 0 & \cdots & 0 & 0 & h_0 \end{bmatrix} \begin{bmatrix} 0 \\ W_{1,0} \\ 0 \\ W_{1,1} \\ 0 \\ W_{1,2} \\ \vdots \\ W_{1,\frac{N}{2}-2} \\ 0 \\ W_{1,\frac{N}{2}-1} \end{bmatrix}.$$

在 $\mathbf{W}_1$ 中插入零的过程在工程文献中称为上采样(采样因子为 2), 在某种意义上, 上采样对下采样是互补的过程(参见练习[4.9]). 如果我们定义

$$W_{1,t}^{\uparrow} \equiv \begin{cases} 0, & t=0,2,\cdots,N-2; \\ W_{1,\frac{t-1}{2}}, & t=1,3,\cdots,N-1. \end{cases}$$

我们能写成

$$\mathcal{D}_{1,t} = \sum_{l=0}^{L-1} h_l W_{1,t+l \bmod N}^{\uparrow} = \sum_{l=0}^{N-1} h_l^{\circ} W_{1,t+l \bmod N}^{\uparrow}, \quad t=0,1,\cdots,N-1.$$

如果和上面的式(37c)对比一下的话, 我们会发现 $\mathcal{D}_1$ 的元素可以通过 $\mathbf{W}_1$ 上采样后再和滤波器 $\{h_l^{\circ}\}$ 进行互相关运算得到. 由于 $\{h_l^{\circ}\} \leftrightarrow \left\{H\left(\frac{k}{N}\right)\right\}$, 所以可以认为

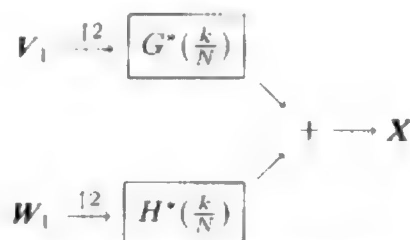


图 83 从 $\boldsymbol{W}_1$ 和 $\boldsymbol{V}_1$ 重构 $\boldsymbol{X}$ 的流程图. 长为 $N/2$ 的向量 $\boldsymbol{W}_1$ 经过一个因子为 2 的上采样(因子为 2 的上采样即在 $\boldsymbol{W}_1$ 的每两个元素中插入一个零)得到一个长为 N 的向量，这个向量再经过滤波器 $\left\{H^*\left(\frac{k}{N}\right)\right\}$ 的循环滤波得到一个输出，类似地，向量 $\boldsymbol{V}_1$ 上采样后经过 $\left\{G^*\left(\frac{k}{N}\right)\right\}$ 的滤波也得到一个输出，将这两个输出相加就得到向量 $\boldsymbol{X}$

$\mathcal{D}_1$ 是 $\boldsymbol{W}_1$ 上采样后通过一个频域表示为 $\left\{H^*\left(\frac{k}{N}\right)\right\}$ 的周期滤波器滤波的结果(参见练习[31]).

类似地，$\mathcal{S}_1$ 的元素可以认为是 $\boldsymbol{V}_1$ 上采样后通过一个频域表示为 $\left\{G^*\left(\frac{k}{N}\right)\right\}$ 的周期滤波器滤波的结果. 如同 $W^{\uparrow}_{1,t}$ 我们可以类似地定义 $V^{\uparrow}_{1,t}$，现在可以写成

$$\begin{aligned} X_t &= \sum_{l=0}^{L-1} h_l W^{\uparrow}_{1,t+l \bmod N} + \sum_{l=0}^{L-1} g_l V^{\uparrow}_{1,t+l \bmod N} \\ &= \sum_{l=0}^{N-1} h^{\circ}_l W^{\uparrow}_{1,t+l \bmod N} + \sum_{l=0}^{N-1} g^{\circ}_l V^{\uparrow}_{1,t+l \bmod N}, \quad t = 0,1,\cdots, N-1. \end{aligned} \tag{83}$$

图 83 展示了从 $\boldsymbol{W}_1$ 和 $\boldsymbol{V}_1$ 重构 $\boldsymbol{X}$ 的流程图.

现在我们考虑序列 $\boldsymbol{X}$(塔式算法中第一步的输入序列)的离散傅里叶变换及 $\boldsymbol{V}_1$ 和 $\boldsymbol{W}_1$(输出序列)的离散傅里叶变换之间的关系. 由于 $\{X_t\}\leftrightarrow\{\mathcal{X}_k\}$，通过离散傅里叶逆变换(式(36f))可得

$$X_t = \frac{1}{N}\sum_{k=0}^{N-1} \mathcal{X}_k e^{i2\pi tk/N} = \frac{1}{N}\sum_{k=-\frac{N}{2}+1}^{\frac{N}{2}} \mathcal{X}_k e^{i2\pi tk/N}, \quad t = 0,\cdots,N-1$$

注意由于 $\{\mathcal{X}_k\}$ 和 $\{e^{i2\pi tk/N}\}$ 是周期为 N 的序列，因此它们的乘积序列 $\{\mathcal{X}_k e^{i2\pi tk/N}\}$ 也是周期为 N 的序列，所以它们在任意相邻 N 个元素的求和都是不变的. 由于 $\{2^{1/2}\widetilde{V}_{1,t}\}$ 是用标称通带为 $[-1/4, 1/4]$ 低通滤波器 $\{g_l\}$ 对序列 $\{X_t\}$ 滤波得到的，又因为 $\mathcal{X}_k$ 是在频率 $f_k \equiv k/N$ 处的频谱，因此

$$\widetilde{V}_{1,t} \approx \frac{1}{N}\sum_{k=-\frac{N}{4}+1}^{\frac{N}{4}} \mathcal{X}_k e^{i2\pi tk/N}, \quad t = 0,\cdots,N-1.$$

现在考虑下采样序列：属于同样的逼近阶

$$V_{1,t} = 2^{1/2}\widetilde{V}_{1,2t+1} \approx \frac{\sqrt{2}}{N}\sum_{k=-\frac{N}{4}+1}^{\frac{N}{4}} \mathcal{X}_k e^{i2\pi(2t+1)k/N}, \quad t = 0,\cdots,\frac{N}{2}-1$$

$$= \frac{2}{N}\sum_{k=-\frac{N}{4}+1}^{\frac{N}{4}} \frac{\mathcal{X}_k e^{i2\pi k/N}}{\sqrt{2}} e^{i2\pi tk/(N/2)}$$

$$= \frac{1}{N'}\sum_{k=-\frac{N'}{2}+1}^{\frac{N'}{2}} \mathcal{X}'_k e^{i2\pi tk/N'}, \quad t=0,\cdots,N'-1$$

其中 $N'\equiv N/2$，$\mathcal{X}'_k\equiv\mathcal{X}_k e^{i2\pi k/N}/\sqrt{2}$. 从序列$\{V_{1,t}\}$下采样得到的近似的傅里叶序列$\{\mathcal{X}'_k\}$的元素是和频率 $f'_k\equiv k/N'$ 相联系的，其范围为

$$从 \quad f'_{-\frac{N'}{2}+1}=-\frac{1}{2}+\frac{1}{N'} \quad 到 \quad f'_{\frac{N'}{2}}=\frac{1}{2}.$$

由于序列$\{\widetilde{V}_{1,t}\}$可以称为是半带序列，因为它是一个通带为$[-1/4, 1/4]$的低通滤波器的输出，因此这个序列没有高频分量，对$\{V_{1,t}\}$下采样后得到的序列是"满带序列"，因为在整个 $f'_k\in[-1/2, 1/2]$的频率范围内都有不可忽略的分量. 注意$\mathcal{X}_k$ 和$\mathcal{X}'_k$是相对应的：$\mathcal{X}_k$ 相应于频率 $f_k=k/N$，而$\mathcal{X}'_k$相应于频率 $f'_k=k/N'=2k/N=2f_k$. 因此序列$\{X_t\}$的傅里叶表示在频率范围$[-1/4, 1/4]$内的 f_k 的系数和$\{V_{1,t}\}$的傅里叶近似表示在频率范围$[-1/2, 1/2]$内的 f'_k的系数是相对应的.

现在我们用同样的方法来研究小波系数. 因为$\{2^{1/2}\widetilde{W}_{1,t}\}$是用标称通带为 $1/4\leqslant|f|\leqslant1/2$ 的高通滤波器$\{h_l\}$对$\{X_t\}$滤波之后形成的，由此得到

$$\widetilde{W}_{1,t}\approx\frac{1}{N}\left(\sum_{k=-\frac{N}{2}+1}^{-\frac{N}{4}}+\sum_{k=\frac{N}{4}+1}^{\frac{N}{2}}\right)\mathcal{X}_k e^{i2\pi tk/N},$$

$t=0, \cdots, N-1$，其中定义

$$\left(\sum_{k=l}^{m}+\sum_{k=l'}^{m'}\right)A_k\equiv\sum_{k=l}^{m}A_k+\sum_{k=l'}^{m'}A_k.$$

对于下采样后得到的序列$\left\{W_{1,t}: t=0, \cdots, \frac{N}{2}-1\right\}$，我们有

$$W_{1,t}=2^{1/2}\widetilde{W}_{1,2t+1}\approx\frac{\sqrt{2}}{N}\left(\sum_{k=-\frac{N}{2}+1}^{-\frac{N}{4}}+\sum_{k=\frac{N}{4}+1}^{\frac{N}{2}}\right)\mathcal{X}_k e^{i2\pi(2t+1)k/N}$$

$$=\frac{2}{N}\left(\sum_{k=-\frac{N}{2}+1}^{-\frac{N}{4}}+\sum_{k=\frac{N}{4}+1}^{\frac{N}{2}}\right)\frac{\mathcal{X}_k e^{i2\pi k/N}}{\sqrt{2}}e^{i2\pi tk/(N/2)}$$

$$=\frac{1}{N'}\sum_{k=-\frac{N'}{2}+1}^{\frac{N'}{2}}\mathcal{X}'_k e^{i2\pi tk/N'},$$

其中现在

$$\mathcal{X}'_k \equiv \frac{\mathcal{X}_{k+\frac{N}{2}} \mathrm{e}^{\mathrm{i}2\pi\left(k+\frac{N}{2}\right)/N}}{\sqrt{2}} = -\frac{\mathcal{X}_{k+\frac{N}{2}} \mathrm{e}^{\mathrm{i}2\pi k/N}}{\sqrt{2}}.$$

和前面一样，子采样序列$\{W_{1,t}\}$的近似傅里叶变换$\{\mathcal{X}'^{k}\}$的元素与频率f'_k相联系，且频率f'_k范围由$-\frac{1}{2}+\frac{1}{N'}$到$\frac{1}{2}$. 因此，$\{\widetilde{W}_{1,t}\}$是一个在低频部分能量分布很小的半带序列(因为它是高通滤波的结果)，下采样后得到的序列$\{W_{1,t}\}$是满带的；即这个序列在所有频率f'_k上的能量都不可忽略. 注意$\mathcal{X}_{k+\frac{N}{2}}$和$\mathcal{X}'_k$是一一对应的，$\mathcal{X}_{k+\frac{N}{2}}$对应频率$f_{k+\frac{N}{2}}=\frac{k}{N}+\frac{1}{2}$；$\mathcal{X}'_k$对应频率$f'_k=k/N'=2k/N=2f_k$. 随着$k$从0变化到$N'/2$，$f'_k$在区间[0，1/2]上变化，而$f_{k+\frac{N}{2}}$从$\frac{1}{2}$变化到$\frac{N'}{2N}+\frac{1}{2}=\frac{3}{4}$，即在区间[1/2，3/4]上变化. 对于实值序列，频率f_k处的傅里叶系数在区间[1/2，3/4]上是与频率$1-f_k$(在区间[1/4，1/2]上)相关的系数的复共轭. 这个从[1/2，3/4]到[1/4，1/2]的映射是逆序的，即当我们在[1/2，3/4]中从左向右扫掠时，在[1/4，1/2]中是从右向左扫掠的. 于是$\{X_t\}$的傅里叶表示的频率$f_k\in[1/4, 1/2]$的系数的复共轭$\mathcal{X}_k^*$和$\{W_{1,t}\}$的傅里叶近似表示的频率$f'_k\in[0, 1/2]$的系数的复共轭$\mathcal{X}'^*_k$是一一映上的(以相反顺序).

塔式算法的第一步从长为N的满带序列$\{X_t\}$开始，将其变换成两个新的满带序列，即第一层尺度系数$\{V_{1,t}\}$和第一层小波系数$\{W_{1,t}\}$，这两个序列的长度都是$N/2$. 尺度系数是序列$\{X_t\}$低频部分的近似，而小波系数是序列$\{X_t\}$高频部分的近似. 图86用长为32的时间序列形象地表示了以上思想. 对$t=0, \cdots, 31$，

$$X_t = \sum_{k=0}^{31} \mathcal{X}_k \mathrm{e}^{\mathrm{i}2\pi tk/32},$$

而

$$\mathcal{X}_k \equiv \begin{cases} 0.96-0.03k, & k=0,\cdots,13 \text{ 或 } 16; \\ 0.84, & k=14; \\ 0.735, & k=15; \\ \mathcal{X}_{32-k}, & k=17,\cdots,31. \end{cases} \tag{85}$$

最后，让我们考虑和尺度系数$\{V_{1,t}\}$相关的尺度的概念. 对于哈尔离散小波变换的特殊情形，$g_0=g_1=1/\sqrt{2}$，我们有

$$V_{1,t} = 2^{1/2}\widetilde{V}_{1,2t+1} = g_0 X_{2t+1} + g_1 X_{2t} = \frac{X_{2t+1}+X_{2t}}{\sqrt{2}} \propto \overline{X}_{2t+1}(2);$$

即，$\{V_{1,t}\}$与$\{X_t\}$在尺度2上的平均成正比. 类似地，对于离散 Daubechies 系的其他有限长度的小波类来说，序列$\{V_{1,t}\}$也都可以认为是与尺度2上的加权平均成正比的(详情参见4.8节). 下面举个例子来说明这一点. D(4)尺度滤波器(选择了适当的权重a和b)相当于用两点滤波器$\{a, b\}$光滑$\{X_t\}$得到序列$\{aX_t+bX_{t-1}\}$，然后用滤波器$\{1/4, 1/2, 1/4\}$对这个新序列进行光滑. 注意$\{V_{1,t}\}$相当于在尺度2上进行平均，而$\{W_{1,t}\}$则是对尺度1上变化率的某种度量. 因此输出的尺度系数和小波系数的尺度总是相差两倍. 在塔式算法中这一点区别非常重要，我们用$\lambda_j\equiv 2^j$表示尺度滤波器输出序列的尺度，而用$\tau_j\equiv 2^{j-1}$表示小波滤波器输出序列的尺度(因此在第一部分解中，$\lambda_1=2$而$\tau_1=1$).

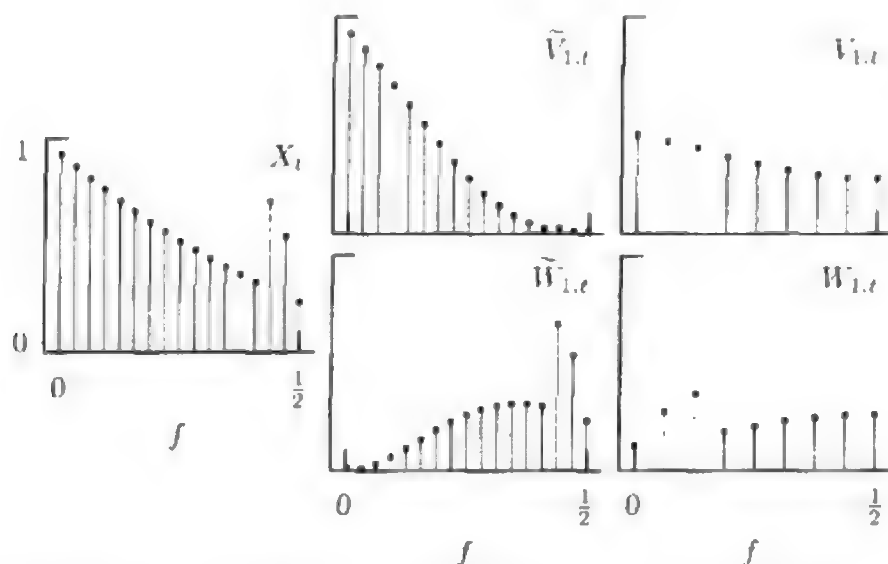


图 86 式(85)中的时间序列$\{X_t\}$的离散傅里叶变换值的平方(左图)，和哈尔小波滤波器和尺度滤波器的输出$\{\widetilde{V}_{1,t}\}$和$\{\widetilde{W}_{1,t}\}$的离散傅里叶变换值的平方(中图)，还有哈尔小波系数和尺度系数$\{V_{1,t}\}$和$\{W_{1,t}\}$(右图)．由于$\{X_t\}$和它的滤波输出是实值的，因此所有的离散傅里叶变换是关于零点对称的，因此图中只显示了非负的傅里叶频率．$\{\widetilde{V}_{1,t}\}$和$\{\widetilde{W}_{1,t}\}$的离散傅里叶变换值的平方函数是通过将$\{X_t\}$的平方幅度函数乘以$\{g_l/\sqrt{2}\}$和$\{h_l/\sqrt{2}\}$的平方增益函数(定义为$\mathcal{G}(f)/2$和$\mathcal{H}(f)/2$——它们的图形在图 73 中给出)后得到的．注意到$\{\widetilde{V}_{1,t}\}$为序列$\{X_t\}$的低频成分，而$\{\widetilde{W}_{1,t}\}$为序列$\{X_t\}$的高频成分．$\{\widetilde{V}_{1,t}\}$在高频部分能量很小，$\{\widetilde{W}_{1,t}\}$在低频部分能量很小．因此这两个序列都是半带序列，下采样后得到的序列$\{V_{1,t}\}$和$\{W_{1,t}\}$都是满带滤波器，$\{V_{1,t}\}$保留了$\{X_{1,t}\}$低频部分，$\{W_{1,t}\}$保留了$\{X_{1,t}\}$高频部分．但是$\{V_{1,t}\}$保持了$\{X_t\}$的频率的顺序，而$\{W_{1,t}\}$的频率和$\{X_t\}$的频率差了一个反序．例如，$\{X_t\}$的离散傅里叶变换的高频成分$f_{14}=\frac{14}{32}$和$f_{15}=\frac{15}{32}$在$\{W_{1,t}\}$的离散傅里叶变换中表现为低频成分$f'_1=\frac{1}{16}$和$f'_2=\frac{2}{16}$

4.4 节的关键结论与定义

塔式算法的第一步是对$\{X_t,\ t=0,\ \cdots,\ N-1\}$进行正交分解，变成两个新的序列，即$\left\{W_{1,t}:\ t=0,\ \cdots,\ \frac{N}{2}-1\right\}$和$\left\{V_{1,t}:\ t=0,\ \cdots,\ \frac{N}{2}-1\right\}$．关于第一层小波系数$\{W_{1,t}\}$有如下结论：

[1]$\mathcal{W}_1$是一个$\frac{N}{2}\times N$阶矩阵，且满足$\mathcal{W}_1\mathcal{W}_1^{\mathrm{T}}=I_{N/2}$，并且由离散小波变换频率矩阵$\mathcal{W}$的前$N/2$行组成$\left(\mathcal{W}_1\text{ 的每一行包含}\{h_l^\circ\}\leftrightarrow H\left(\frac{k}{N}\right)\text{的元素}\right)$．即小波滤波器$\{h_l\}$以$N$为周期进行周期化；第一行为$[h_1^\circ,\ h_0^\circ,\ h_{N-1}^\circ,\ h_{N-2}^\circ,\ \cdots,\ h_2^\circ]$，剩下的$\frac{N}{2}-1$行是由第一行通过右平移得到的，平移长度分别为 2，4，6，…，$N-2$．

[2]向量$\mathbf{W}_1=\mathcal{W}_1\mathbf{X}$的元素个数为$N/2$．

[3]第一层细节$\mathcal{D}_1=\mathcal{W}_1^{\mathrm{T}}\mathbf{W}_1=\mathcal{W}_1^{\mathrm{T}}\mathcal{W}_1\mathbf{X}$可以通过对$\mathbf{W}_1$进行因子为 2 的上采样然后用$\left\{H^*\left(\frac{k}{N}\right)\right\}$进行滤波得到．

[4]$\{W_{1,t}\}$表现的是尺度$\tau_1=1$的变化率．

[5]$\{W_{1,t}\}$是对时间序列$\{X_t\}$近似在[1/4，1/2]内的高通滤波结果再通过因子为 2 的下采样得到的.

[6]$\{W_{1,t}\}$对应$\{X_t\}$的离散傅里叶变换在[1/4，1/2]内的那部分，只是频率上顺序颠倒了.

对于第一层尺度系数则有如下结论：

[1]$\mathcal{V}_1$ 是一个$\frac{N}{2}\times N$ 阶矩阵，且满足$\mathcal{V}_1\mathcal{V}_1^{\mathrm{T}}=I_{N/2}$，$\mathcal{W}_1\mathcal{V}_1^{\mathrm{T}}=\mathcal{V}_1\mathcal{W}_1^{\mathrm{T}}=0_{N/2}$，且各行和离散傅里叶变换矩阵$\mathcal{W}$的后 $N/2$ 行张成相同的 $\mathbf{R}^N$ 的子空间.（$\mathcal{V}_1$ 和$\mathcal{W}_1$ 有相同的结构，只需将 h_l°换成 g_l°就行了，其中$\{g_l^\circ\}$是尺度滤波器$\{g_l\}$以 N 为周期进行周期化的结果，即有$\{g_l^\circ\}\leftrightarrow\left\{G\left(\frac{k}{N}\right)\right\}$.

[2]$\mathbf{V}_1=\mathcal{V}_1\boldsymbol{X}$ 的元素个数为 $N/2$.

[3]$\{V_{1,t}\}$的第一层光滑 $\mathcal{S}_1=\mathcal{V}_1^{\mathrm{T}}\mathbf{V}_1=\mathcal{V}_1^{\mathrm{T}}\mathcal{V}_1\boldsymbol{X}$，可以通过对 $\mathbf{V}_1$ 进行因子为 2 的上采样然后再用$\left\{G^*\left(\frac{k}{N}\right)\right\}$进行滤波得到.

[4]$\{V_{1,t}\}$表现的是对时间序列以尺度 $\lambda_1=2$ 进行平均.

[5]$\{V_{1,t}\}$是对时间序列$\{X_t\}$近似在[0，1/4]内的低通滤波结果再通过因子为 2 的下采样得到的.

[6]$\{V_{1,t}\}$是$\{X_t\}$的离散傅里叶变换在频率[0，1/4]内的部分.

序列$\{W_{1,t}\}$组成小波系数 $\boldsymbol{W}$ 向量的前半部分；余下的一半由$\{V_{1,t}\}$在塔式算法相继的步骤中得到.

4.4 节的评论与扩展

[1]如果我们对用离散傅里叶变换的方法做一些修改，直接将采样间隔 Δt 考虑进去，就会对$\{X_t\}$和$\{V_{1,t}\}$的频域关系有更深刻的认识. 现在有 $f_k=k/(N\Delta t)$，Δt 如果是 1 秒的话，f_k 就是每隔一秒得到的采样数据. 考虑了 Δt，$\{X_t\}$的离散傅里叶变换和离散傅里叶逆变换就成了如下形式：

$$\mathcal{X}_k=\Delta t\sum_{t=0}^{N-1}X_t\mathrm{e}^{-\mathrm{i}2\pi f_k t\Delta t},\quad X_t=\frac{1}{N\Delta t}\sum_{k=-\frac{N}{2}+1}^{\frac{N}{2}}\mathcal{X}_k\mathrm{e}^{\mathrm{i}2\pi f_k t\Delta t}.$$

这里频率满足$-f_{\mathcal{N}}<f_k\leqslant f_{\mathcal{N}}$，其中 $f_{\mathcal{N}}\equiv1/(2\Delta t)$称作是奈奎斯特频率. 用$\{g_l\}$对时间序列循环滤波的结果可以近似地表示成

$$\widetilde{V}_{1,t}\approx\frac{1}{N\Delta t}\sum_{k=-\frac{N}{4}+1}^{\frac{N}{4}}\mathcal{X}_k\mathrm{e}^{\mathrm{i}2\pi f_k t\Delta t},$$

频率范围满足$-f_{\mathcal{N}}/2<f_k\leqslant f_{\mathcal{N}}/2$. 由于对 $\widetilde{V}_{1,t}$的采样间隔和 X_t 相同，所以下采样序列的采样间隔为 $2\Delta t\equiv\Delta t'$. 因为$\{V_{1,t}\}$的长度为 $N'\equiv N/2$，所以它的离散傅里叶变换的傅里叶频率为 $k/(N'\Delta t')=k/(N\Delta t)=f_k$，其中 k 的取值范围现在为从$-\frac{N}{4}+1$ 到$\frac{N}{4}$. 因此$\{X_t\}$和$\{V_{1,t}\}$的傅里叶变换形式是相同的，但是前一种情况

满足$-f_N<f_k\leqslant f_N$，而第二种情况则满足$-f_N/2<f_k\leqslant f_N/2$(注意$f_N/2=1/(2\Delta t')$，正好是下采样序列的奈奎斯特频率)．现在我们能写成：

$$V_{1,t}\approx\frac{1}{N'\Delta t'}\sum_{k=-\frac{N'}{2}+1}^{\frac{N'}{2}}\mathcal{X}'_k e^{i2\pi f_k t\Delta t'},\quad 其中现在\mathcal{X}'_k=\sqrt{2}\,\mathcal{X}_k e^{i2\pi f_k\Delta t}.$$

使用包含物理意义下单位的频率的离散傅里叶变换，我们发现$\{X_t\}$的傅里叶表示的频率$f_k\in[-f_N/2,f_N/2]$和$\{V_{1,t}\}$的近似傅里叶表示的频率范围是一样的．将$\{X_t\}$和$\{V_{1,T}\}$表示的物理频率映射到标准的频率区间$[-1/2,1/2]$上，就得到了我们在这一节前面部分描述的方式了(类似地，我们可以证明$\{X_t\}$的傅里叶表示频率$f_k\in[f_N/2,f_N]$以相反顺序映射到$\{W_{1,t}\}$的近似傅里叶表示的频率$f_k\in[0,f_N/2]$上)．

4.5　塔式算法的第二步

就像我们在前面一节开始所描述的那样，塔式算法的第二步就是将$\{V_{1,t}\}$看成是第一步中的$\{X_t\}$．直观上看这样做是很自然的，若将$\{X_t\}$看成是单位尺度上的平均，$\{V_{1,t}\}$是在尺度2上的平均，$\{X_t\}$和$\{V_{1,t}\}$是在不同尺度上的平均，它们应该是很相似的．因此，我们对$\{V_{1,t}\}$用$\{h_l\}$和$\{g_l\}$进行循环滤波并下采样得到两个序列，即

$$W_{2,t}\equiv\sum_{l=0}^{L-1}h_l V_{1,2t+1-l\bmod\frac{N}{2}},\quad V_{2,t}\equiv\sum_{l=0}^{L-1}g_l V_{1,2t+1-l\bmod\frac{N}{2}},\tag{88}$$

$t=0,\cdots,\frac{N}{4}-1$．注意上式和式(70c)及(77b)滤波形式是相同的，只是将“mod$N$”换成了“mod $\frac{N}{2}$”．对尺度2及$j=2$层上的离散小波变换的小波系数用$\mathbf{W}_{2,t}$来记；即

$$\begin{aligned}\mathbf{W}_2&\equiv[W_{\frac{N}{2}},W_{\frac{N}{2}+1},\cdots,W_{\frac{3N}{4}-1}]^{\mathrm{T}}\\&=[W_{2,0},W_{2,1},\cdots,W_{2,\frac{N}{4}-1}]^{\mathrm{T}}.\end{aligned}$$

定义$j=2$层时包含尺度系数的向量为

$$\mathbf{V}_2\equiv[V_{2,0},V_{2,1},\cdots,V_{2,\frac{N}{4}-1}]^{\mathrm{T}}.$$

令$\mathcal{B}_2$和$\mathcal{A}_2$是$\frac{N}{4}\times\frac{N}{2}$阶矩阵，其行向量分别包含了$\{h_l\}$和$\{g_l\}$的周期为$N/2$的循环平移，我们可以将$\mathbf{V}_1$变换到$\mathbf{W}_2$和$\mathbf{V}_2$：

$$\begin{bmatrix}\mathbf{W}_2\\\mathbf{V}_2\end{bmatrix}=\mathcal{P}_2\mathbf{V}_1=\begin{bmatrix}\mathcal{B}_2\\\mathcal{A}_2\end{bmatrix}\mathbf{V}_1,\quad 其中\mathcal{P}_2\equiv\begin{bmatrix}\mathcal{B}_2\\\mathcal{A}_2\end{bmatrix}.$$

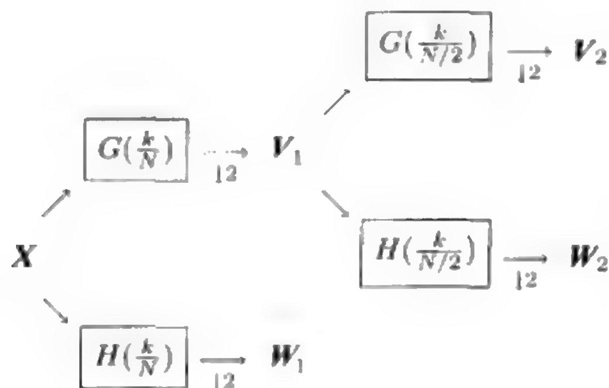


图 89 将 $\boldsymbol{X}$ 分解为 $\boldsymbol{W}_1$，$\boldsymbol{W}_2$ 和 $\boldsymbol{V}_2$ 的流程图

类似$\mathcal{W}_1$ 和$\mathcal{V}_1$ 的情形，$\mathcal{B}_2$ 和$\mathcal{A}_2$ 的各行也是规范正交的：

$$\mathcal{B}_2\mathcal{A}_2^{\mathrm{T}} = \mathcal{A}_2\mathcal{B}_2^{\mathrm{T}} = 0_{\frac{N}{4}}, \quad \mathcal{B}_2\mathcal{B}_2^{\mathrm{T}} = \mathcal{A}_2\mathcal{A}_2^{\mathrm{T}} = I_{\frac{N}{4}}.$$

因此$\frac{N}{2}\times\frac{N}{2}$阶矩阵$\mathcal{P}_2$ 是规范正交的，所以我们可以通过下式重构 $\boldsymbol{V}_1$：

$$\boldsymbol{V}_1 = \mathcal{P}_2^{\mathrm{T}}\begin{bmatrix}\boldsymbol{W}_2\\ \\ \boldsymbol{V}_2\end{bmatrix} = [\mathcal{B}_2^{\mathrm{T}} \quad \mathcal{A}_2^{\mathrm{T}}]\begin{bmatrix}\boldsymbol{W}_2\\ \\ \boldsymbol{V}_2\end{bmatrix} = \mathcal{B}_2^{\mathrm{T}}\boldsymbol{W}_2 + \mathcal{A}_2^{\mathrm{T}}\boldsymbol{V}_2.$$

将上式代入式(81a)得

$$\boldsymbol{X} = \mathcal{W}_1^{\mathrm{T}}\boldsymbol{W}_1 + \mathcal{V}_1^{\mathrm{T}}\ \mathcal{B}_2^{\mathrm{T}}\boldsymbol{W}_2 + \mathcal{V}_1^{\mathrm{T}}\ \mathcal{A}_2^{\mathrm{T}}\boldsymbol{V}_2 = \mathcal{B}_1^{\mathrm{T}}\boldsymbol{W}_1 + \mathcal{A}_1^{\mathrm{T}}\mathcal{B}_2^{\mathrm{T}}\boldsymbol{W}_2 + \mathcal{A}_1^{\mathrm{T}}\mathcal{A}_2^{\mathrm{T}}\boldsymbol{V}_2,$$

为了以后描述塔式算法方便起见，令$\mathcal{A}_1\equiv\mathcal{V}_1$，$\mathcal{B}_1\equiv\mathcal{W}_1$. 对比式(63c)，我们可以得到$\mathcal{W}_2\equiv\mathcal{B}_2\mathcal{A}_1$，于是

$$\mathcal{D}_2 = \mathcal{W}_2^{\mathrm{T}}\boldsymbol{W}_2 = \mathcal{A}_1^{\mathrm{T}}\mathcal{B}_2^{\mathrm{T}}\boldsymbol{W}_2.$$

由于$\mathcal{D}_1=\mathcal{B}_1^{\mathrm{T}}\boldsymbol{W}_1$，我们得到如下两式：

$$\boldsymbol{X} = \mathcal{D}_1 + \mathcal{D}_2 + \mathcal{A}_1^{\mathrm{T}}\mathcal{A}_2^{\mathrm{T}}\boldsymbol{V}_2 \quad \text{和} \quad \boldsymbol{X} = \mathcal{D}_1 + \mathcal{D}_2 + \mathcal{S}_2,$$

由此得到

$$\mathcal{S}_2 = \mathcal{A}_1^{\mathrm{T}}\mathcal{A}_2^{\mathrm{T}}\boldsymbol{V}_2 \equiv \mathcal{V}_2^{\mathrm{T}}\boldsymbol{V}_2,$$

其中$\mathcal{V}_2\equiv\mathcal{A}_2\mathcal{A}_1$. 图 89 与 90 画的流程图就是用塔式算法将 $\boldsymbol{X}$ 分解成 $\boldsymbol{W}_1$，$\boldsymbol{W}_2$ 和 $\boldsymbol{V}_2$ 及用塔式算法从 $\boldsymbol{W}_1$，$\boldsymbol{W}_2$ 和 $\boldsymbol{V}_2$ 重构 $\boldsymbol{X}$ 的过程.

因为序列$\{W_{2,t}\}$和$\{V_{2,t}\}$是对$\{V_{1,t}\}$滤波之后再下采样得到的，又由于$\{V_{1,t}\}$也是对$\{X_t\}$滤波之后再下采样得到的，因此我们可以用包含下采样的级联算法来描述如何从$\{X_t\}$得到$\{W_{2,t}\}$和$\{V_{2,t}\}$. 为此定义$\{h_l^{\uparrow}\}$是宽为 $2L-1$ 的具有系数

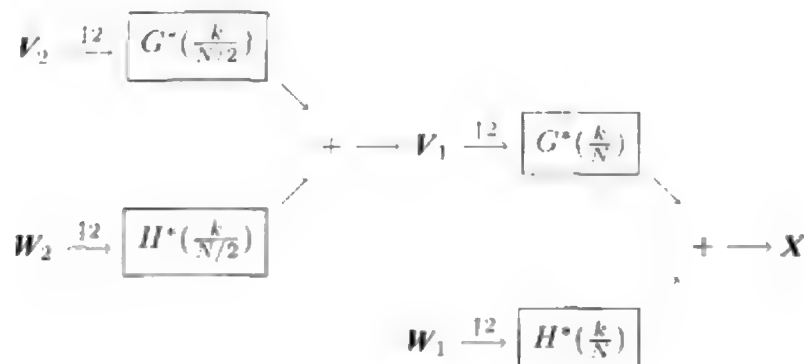


图 90　从 $\boldsymbol{W}_1$，$\boldsymbol{W}_2$ 和 $\boldsymbol{V}_2$ 重构 $\boldsymbol{X}$ 的流程图

h_0，0，h_1，0，…，h_{L-2}，0，h_{L-1}的滤波器，即通过在$\{h_l\}$的 L 的元素中插入零来得到$\{h_l^{\uparrow}\}$．令

$$2\widetilde{W}_{2,t} \equiv \sum_{l=0}^{2L-2} h_l^{\uparrow}\, 2^{1/2}\widetilde{V}_{1,t-l \bmod N}, \quad t = 0,\cdots,N-1. \tag{90a}$$

式(77a)告诉我们$\{2^{1/2}\widetilde{V}_{1,t}\}$是用$\{g_l\}$对$\{X_t\}$进行滤波得到的．因此$\{2\widetilde{W}_{2,t}\}$是级联滤波器的输出，即是对$\{X_t\}$先用$\{g_l\}$滤波再用$\{h_l^{\uparrow}\}$滤波的结果．我们要求

$$W_{2,t} = 2\widetilde{W}_{2,4t+3}, \quad t = 0,\cdots,\frac{N}{4}-1, \tag{90b}$$

即$\{W_{2,t}\}$可以从$\{2\widetilde{W}_{2,t}\}$中每四个抽取一个得到，第一个抽取的元素为 $2\widetilde{W}_{2,3}$（用工程的术语来说这就是对$\{2\widetilde{W}_{2,t}\}$进行因子为 4 的下采样）．为了证实上述说法，注意有

$$\begin{aligned}
2\widetilde{W}_{2,4t-3} &= \sum_{l=0}^{2L-2} h_l^{\uparrow}\, 2^{1/2}\widetilde{V}_{1,4t+3-l \bmod N} \\
&= \sum_{l=0}^{L-1} h_l 2^{1/2}\widetilde{V}_{1,4t+3-2l \bmod N} \\
&= \sum_{l=0}^{L-1} h_l 2^{1/2}\widetilde{V}_{1,2(2t+1-l)+1 \bmod N} \\
&= \sum_{l=0}^{L-1} h_l V_{1,2t+1-l \bmod \frac{N}{2}} = W_{2,t},
\end{aligned}$$

推导过程中我们用了式(77b)和(88)．令$\{h_{2,l}\} \equiv \{g * h_l^{\uparrow}\}$是$\{g_l\}$与$\{h_l^{\uparrow}\}$的卷积．由于$\{g_l\}$的长度为 L 而$\{h_l^{\uparrow}\}$的长度为 $2L-1$，因此用练习[28a]的结论可得$\{h_{2,l}\}$的长度为：

$$L_2 \equiv L + (2L-1) - 1 = 3L - 2.$$

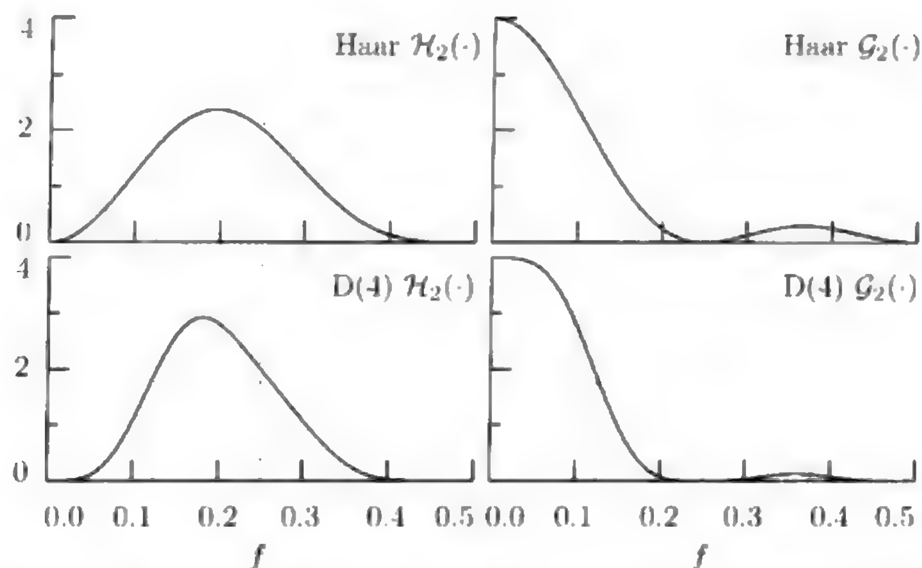


图 91 哈尔滤波器 $\{h_{2,l}\}$ 和 $\{g_{2,l}\}$ 的平方增益函数(分别为左上图和右上图)及 D(4) 滤波器 $\{h_{2,l}\}$ 和 $\{g_{2,l}\}$ 的平方增益函数(分别为左下图和右下图). 图中用点划线标记出了频率 $f=1/8$ 和 $f=1/4$

练习[91] 设 $\{h_l: l=0, \cdots, L-1\}$ 是一个滤波器，传递函数为 $H(\cdot)$. 通过在 $\{h_l\}$ 的元素中插入 m 个零定义新的滤波器为

$$h_0, \underbrace{0, \cdots, 0}_{m\text{个}0}, h_1, \cdots, h_{L-2}, \underbrace{0, \cdots, 0}_{m\text{个}0}, h_{L-1}.$$

证明这个滤波器的传递函数是 $H([m+1]f)$. ◁

令上面的 $m=1$，我们可以得到 $\{h_{2,l}\}$ 的传递函数由下式给出：

$$H_2(f) \equiv H(2f)G(f).$$

用类似的方法，得到 $\{g_{2,l}\}$ 是对 $\{X_t\}$ 滤波后得到 $\{V_{2,t}\}$ 的滤波器，其宽度为 L_2，通过将 $\{g_l\}$ 和

$$\{g_l^{\uparrow}\} \equiv \{g_0, 0, g_1, 0, \cdots, g_{L-2}, 0, g_{L-1}\}$$

进行卷积得到，其传递函数为

$$G_2(f) \equiv G(2f)G(f).$$

图 91 给出了哈尔滤波器和 D(4)滤波器作为 f 函数的

$$\mathcal{H}_2(f) \equiv |H_2(f)|^2 \quad \text{和} \quad \mathcal{G}_2(f) \equiv |G_2(f)|^2$$

的图形. 注意 $\{h_{2,l}\}$ 是一个带通滤波器且其近似通带为 $1/8 \leqslant |f| \leqslant 1/4$，$\{g_{2,l}\}$ 是一个低通滤波器其通带近似为 $0 \leqslant |f| \leqslant 1/8$. 图 92 给出了直接从 $\mathbf{X}$ 计算 $\mathbf{W}_2$ 和 $\mathbf{V}_2$ 的流程图.

为了方便起见，我们称 $\{h_{2,l}\}$ 和 $\{g_{2,l}\}$ 是第二层小波滤波器和尺度滤波器，因为它们可以产生第二层小波系数和尺度系数. 在下一节我们将这个概念推广到第 j 层小波滤波器和尺度滤波器. 我们可以定义第一层小波滤波器 $\{h_{1,l}\}$ 和尺度滤波器 $\{g_{1,l}\}$ 就等于 $\{h_l\}$ 和 $\{g_l\}$.

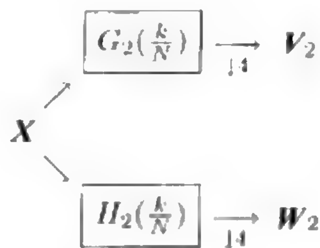


图 92 从 $\boldsymbol{X}$ 直接算出 $\boldsymbol{W}_2$ 和 $\boldsymbol{V}_2$ 的流程图. 长为 N 的时间序列向量 $\boldsymbol{X}$ 用以 N 为周期进行周期化后得到的 $j=2$ 层的小波滤波器 $\{h_{2,l}\}$(其传递函数为 $H_2(\cdot)$)进行循环滤波. 然后保留所有的指标 t 满足 $t+1$ 为 4 的倍数的系数，得到向量 $\boldsymbol{W}_2$("↓4"表示每 4 个系数保留一个的下采样). $\boldsymbol{W}_2$ 的 $N/4$ 的系数是 $j=2$ 层上的小波系数. 类似地，长为 $N/4$ 的向量 $\boldsymbol{V}_2$ 包含了 $j=2$ 层上的尺度系数，它是对 $\boldsymbol{X}$ 用频域表示为 $\left\{G_2\left(\frac{k}{N}\right)\right\}$ 的尺度滤波器进行循环滤波再进行因子为 4 的下采样得到的

练习[92] 证明$\{h_{2,l}\}$求和为零且有单位能量，但它不满足小波滤波器定义中的最后一个性质，即与自身偶平移的规范正交性质(69c). 通过一些特殊的哈尔小波的第二层小波滤波器的例子说明这一点. 提示：考虑小波滤波器的规范正交性质的频域表达式. ◁

如上述练习所表明的，第二层小波滤波器事实上不必是小波滤波器！类似的观察对于第二层尺度滤波器也成立(对于将在下一节定义的 $j\geqslant 3$ 层小波滤波器与尺度滤波器也成立).

因为 $\boldsymbol{W}_2$ 的元素是用$\{h_{2,l}: l=0, \cdots, L_2-1\}$对$\{X_t\}$进行循环滤波得到的，因此也可以通过用$\{h_{2,l}\}$的周期化后的滤波器$\{h^{\circ}_{2,l}: l=0, \cdots, N-1\}$对$\{X_t\}$进行滤波得到. 即我们有

$$W_{2,t}=\sum_{l=0}^{N-1} h^{\circ}_{2,l} X_{4(t+1)-1-l \bmod N}, \quad t=0,\cdots,\frac{N}{4}-1$$

(参见练习[4.10]). 因为我们还有 $\boldsymbol{W}_2=\mathcal{W}_2\boldsymbol{X}$，$\mathcal{W}_2$ 的行一定包含$\{h^{\circ}_{2,l}\}$的循环平移变型. 如果我们注意到 $\boldsymbol{W}_2$ 的第 0 个元素是

$$W_{2,0}=\sum_{l=0}^{N-1} h^{\circ}_{2,l} X_{3-l \bmod N},$$

很明显$\mathcal{W}_2$ 的第一行(或等价地$\mathcal{W}$的第 $N/2$ 行)由下式给出：

$$[h^{\circ}_{2,3}, h^{\circ}_{2,2}, h^{\circ}_{2,1}, h^{\circ}_{2,0}, h^{\circ}_{2,N-1}, h^{\circ}_{2,N-2}, \cdots, h^{\circ}_{2,5}, h^{\circ}_{2,4}]=\mathcal{W}^{\mathrm{T}}_{\frac{N}{2}\cdot}; \tag{92}$$

剩下的$\frac{N}{4}-1$ 行为$\mathcal{T}^{4k}\mathcal{W}^{\mathrm{T}}_{\frac{N}{2}\cdot}$, $k=1, \cdots, \frac{N}{4}-1$，即通过对第一行进行跨度为 4 的循环平移得到. 同样，我们从练习[33]知道$\{h^{\circ}_{2,l}\}$的离散傅里叶变换是对$\{h_{2,l}\}$的传递函数在 $f_k=k/N(k=0, \cdots, N-1)$处进行采样得到. 由于它的传递函数定义为 $H(2f)G(f)$，我们有$\{h^{\circ}_{2,l}\}\leftrightarrow\left\{H\left(\frac{k}{N/2}\right)G\left(\frac{k}{N}\right)\right\}$. 用类似方法，周期滤波器$\{g^{\circ}_{2,l}\}\leftrightarrow\left\{G\left(\frac{k}{N/2}\right)G\left(\frac{k}{N}\right)\right\}$，这个序列生成了$\mathcal{V}_2$ 的各个行向量.

4.5 节中的关键结论与定义

塔式算法的第二步将$\mathcal{V}_1$中的$N/2$个基向量转换成两组有$N/4$个基向量的矩阵，即$\mathcal{W}_2=\mathcal{B}_2\mathcal{V}_1=\mathcal{B}_2\mathcal{A}_1$和$\mathcal{V}_2=\mathcal{A}_2\mathcal{V}_1=\mathcal{A}_2\mathcal{A}_1$，其中我们定义了$\mathcal{A}_1=\mathcal{V}_1$．$\mathcal{B}_2$的各行是由小波滤波器$\{h_l\}$以$N/2$为周期进行周期化$\left(即对\left\{H\left(\frac{k}{N/2}\right):k=0,\cdots,\frac{N}{2}-1\right\}进行离散傅里叶逆变换\right)$且相邻行相差一个跨度为 2 的平移．类似地，$\mathcal{A}_2$ 的各行是由尺度滤波器$\{h_l\}$以 $N/2$ 为周期进行周期化得到的．这一步将$\left\{V_{1,t}:\ t=0,\ \cdots,\ \frac{N}{2}-1\right\}$分解成两个新的序列，即$\left\{W_{2,t}:t=0,\cdots,\frac{N}{4}-1\right\}$和$\left\{V_{2,t}:t=0,\cdots,\frac{N}{4}-1\right\}$．就像第一步通过小波滤波器和尺度滤波器将 $\boldsymbol{X}$ 正交分解成高频成分$\{W_{1,t}\}$和低频成分$\{V_{1,t}\}$，第二步也将$\{V_{1,t}\}$分解成$\{W_{2,t}\}$和$\{V_{2,t}\}$．$\boldsymbol{X}$ 的低频部分$\{V_{1,t}\}$被第二层小波滤波器和尺度滤波器分解成了两部分，这两个滤波器的通带分别为$1/8\leqslant|f|\leqslant1/4$(相应尺度 2 的小波系数)和$0\leqslant|f|\leqslant1/8$．因此关于第二层小波系数$\{W_{2,t}\}$有如下事实：

[1]$\mathcal{W}_2=\mathcal{B}_2\mathcal{A}_1$ 是一个$\frac{N}{4}\times N$阶矩阵且满足$\mathcal{W}_2\mathcal{W}_2^{\mathrm{T}}=I_{N/4}$，构成了离散小波变换矩阵$\mathcal{W}$的从$\frac{N}{2}$到$\frac{3N}{4}-1$的各行$\left(\mathcal{W}_2 的每一行包含元素\{h_{2,l}^{\circ}\}\leftrightarrow\left\{H\left(\frac{k}{N/2}\right)G\left(\frac{k}{N}\right)\right\}，即第二层小波滤波器\{h_{2,l}\}以长度 N 为周期进行周期化得到；第一行向量由式(92)给出，剩下的\frac{N}{4}-1 行是通过循环平移第一行得到的，平移跨度为 4，8，12，\cdots，N-4\right)$．

[2]向量$\boldsymbol{W}_2=\mathcal{W}_2\boldsymbol{X}$的元素长为$N/4$．

[3]第二层细节$\mathcal{D}_2=\mathcal{W}_2^{\mathrm{T}}\boldsymbol{W}_2=\mathcal{W}_2^{\mathrm{T}}\mathcal{W}_2\boldsymbol{X}$．

[4]尺度为$\tau_2=2$的平均变化率．

[5]对第一层尺度系数$\{V_{1,t}\}$高频在[1/4，1/2]半通带进行因子为 2 的下采样得到的结果．

[6]$\{X_t\}$的离散傅里叶变换在频率范围[1/8，1/4]内的部分，只是频率顺序相反．

类似地，第二层尺度系数序列$\{V_{2,t}\}$有如下事实：

[1]$\mathcal{V}_2=\mathcal{A}_2\mathcal{A}_1$ 是一个$\frac{N}{4}\times N$阶矩阵且满足$\mathcal{V}_2\mathcal{V}_2^{\mathrm{T}}=I_{N/4}$，以及$\mathcal{W}_2\mathcal{V}_2^{\mathrm{T}}=\mathcal{V}_2\mathcal{W}_2^{\mathrm{T}}=0_{N/4}$，其各行和离散小波变换矩阵的后 $N/4$ 行向量张成 $\mathbf{R}^N$ 中同一个子空间$\left(\mathcal{V}_2 的结构和\mathcal{W}_2 类似：只要将 h_{2,l}^{\circ}换成 g_{2,l}^{\circ}就行了，其中\{g_{2,l}^{\circ}\}是\{g_{2,l}\}以 N 为周期周期化的结果，即\{g_{2,l}^{\circ}\}\leftrightarrow\left\{G\left(\frac{k}{N/2}\right)G\left(\frac{k}{N}\right)\right\}\right)$．

[2]长为$N/4$的向量$\boldsymbol{V}_2=\mathcal{V}_2\boldsymbol{X}$的元素．

[3]第二层光滑$\mathcal{S}_2=\mathcal{V}_2^{\mathrm{T}}\boldsymbol{V}_2=\mathcal{V}_2^{\mathrm{T}}\mathcal{V}_2\boldsymbol{X}$．

[4]在尺度$\lambda_2=4$上的平均．

[5]对第一层尺度系数$\{V_{1,t}\}$频率在[0，1/4]部分进行因子为 2 的下采样的结果．

[6]$\{X_t\}$的离散傅里叶变换在频率范围[0，1/8]内的部分．

4.6 塔式算法的一般步骤

通过前面两节的讨论，我们现在讨论塔式算法的一般步骤，其中 $j=1,\ \cdots,\ J$

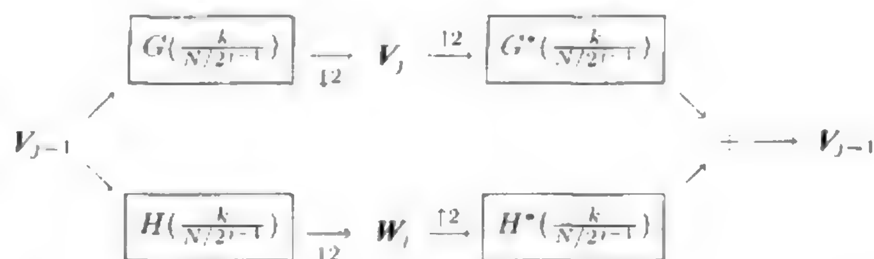


图 94　将 $\mathbf{V}_{j-1}$ 分解成 $\mathbf{W}_j$ 和 $\mathbf{V}_j$，然后再用 $\mathbf{W}_j$ 和 $\mathbf{V}_j$ 合成 $\mathbf{V}_{j-1}$ 的流程图

(前面几节都假定样本长度 N 为 2^J)．定义 $V_{0,t}$ 为 X_t，第 j 步输入就是 $\{V_{j-1,t}: t=0, \cdots, N_{j-1}-1\}$，其中 $N_j \equiv N/2^j$．第 j 步输入可以看成是在尺度 $\lambda_{j-1}=2^{j-1}$ 上尺度系数的平均．第 j 步输出是第 j 层小波系数和尺度系数：

$$W_{j,t} \equiv \sum_{l=0}^{L-1} h_l V_{j-1,2t+1-l \bmod N_{j-1}}, \quad V_{j,t} \equiv \sum_{l=0}^{L-1} g_l V_{j-1,2t+1-l \bmod N_{j-1}}, \tag{94a}$$

$t=0, \cdots, N_j-1$．尺度 $\tau_j=2^{j-1}$ 上的小波系数是 $W_{j,t}$，即

$$\mathbf{W}_j \equiv [W_{N-N_{j-1}}, W_{N-N_{j-1}+1}, \cdots, W_{N-N_j-1}]^{\mathrm{T}} = [W_{j,0}, W_{j,1}, \cdots, W_{j,N_j-1}]^{\mathrm{T}}.$$

如果定义 $\mathbf{V}_j^{\mathrm{T}}$ 的元素为 $V_{j,t}$，那么从 $\mathbf{V}_{j-1}$ 分解成 $\mathbf{W}_j$ 和 $\mathbf{V}_j$，可以写成

$$\begin{bmatrix} \mathbf{W}_j \\ \mathbf{V}_j \end{bmatrix} = \mathcal{P}_j \mathbf{V}_{j-1} = \begin{bmatrix} \mathcal{B}_j \\ \mathcal{A}_j \end{bmatrix} \mathbf{V}_{j-1}, \quad \text{其中} \mathcal{P}_j \equiv \begin{bmatrix} \mathcal{B}_j \\ \mathcal{A}_j \end{bmatrix}, \tag{94b}$$

$\mathcal{B}_j$ 和 $\mathcal{A}_j$ 是 $N_j \times N_{j-1}$ 阶矩阵，其行是由小波滤波器 $\{h_l\}$ 和尺度滤波器 $\{g_l\}$ 以 N_{j-1} 为周期进行周期化得到的．$N_{j-1} \times N_{j-1}$ 阶矩阵 $\mathcal{P}_j$ 是规范正交的，即

$$\mathcal{P}_j \mathcal{P}_j^{\mathrm{T}} = \begin{bmatrix} \mathcal{B}_j \\ \mathcal{A}_j \end{bmatrix} [\mathcal{B}_j^{\mathrm{T}} \quad \mathcal{A}_j^{\mathrm{T}}] = \begin{bmatrix} \mathcal{B}_j \mathcal{B}_j^{\mathrm{T}} & \mathcal{B}_j \mathcal{A}_j^{\mathrm{T}} \\ \mathcal{A}_j \mathcal{B}_j^{\mathrm{T}} & \mathcal{A}_j \mathcal{A}_j^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} I_{N_j} & 0_{N_j} \\ 0_{N_j} & I_{N_j} \end{bmatrix} = I_{N_{j-1}}$$

且

$$\mathcal{P}_j^{\mathrm{T}} \mathcal{P}_j = [\mathcal{B}_j^{\mathrm{T}} \quad \mathcal{A}_j^{\mathrm{T}}] \begin{bmatrix} \mathcal{B}_j \\ \mathcal{A}_j \end{bmatrix} = \mathcal{B}_j^{\mathrm{T}} \mathcal{B}_j + \mathcal{A}_j^{\mathrm{T}} \mathcal{A}_j = I_{N_{j-1}}.$$

如果回忆 $\mathbf{V}_1 = \mathcal{A}_1 \mathbf{X}$，反复用公式 $\mathbf{V}_j = \mathcal{A}_j \mathbf{V}_{j-1}$ 可得

$$\mathbf{W}_j = \mathcal{B}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1 \mathbf{X} = \mathcal{W}_j \mathbf{X}, \quad \text{且} \mathbf{V}_j = \mathcal{A}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1 \mathbf{X} = \mathcal{V}_j \mathbf{X},$$

其中

$$\mathcal{W}_j \equiv \mathcal{B}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1, \quad \mathcal{V}_j \equiv \mathcal{A}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1.$$

由于 $N_{j-1}\times N_{j-1}$ 阶矩阵$\mathcal{P}_j$ 是规范正交的，所以可以用下式重构 $\mathbf{V}_{j-1}$：

$$\mathbf{V}_{j-1}=\mathcal{P}_j^{\mathrm{T}}\begin{bmatrix}\mathbf{W}_j\\ \mathbf{V}_j\end{bmatrix}=[\mathcal{B}_j^{\mathrm{T}}\quad \mathcal{A}_j^{\mathrm{T}}]\begin{bmatrix}\mathbf{W}_j\\ \mathbf{V}_j\end{bmatrix}=\mathcal{B}_j^{\mathrm{T}}\mathbf{W}_j+\mathcal{A}_j^{\mathrm{T}}\mathbf{V}_j. \tag{95a}$$

借助于滤波运算，$\mathbf{V}_{j-1}$的第 t 个元素可以通过类似式(83)重构，即

$$V_{j-1,t}=\sum_{l=0}^{L-1}h_l W^{\uparrow}_{j,t+l \bmod N_{j-1}}+\sum_{l=0}^{L-1}g_l V^{\uparrow}_{j,t+l \bmod N_{j-1}}, \tag{95b}$$

$t=0,1,\cdots,N_{j-1}-1$，其中

$$W^{\uparrow}_{j,t}\equiv\begin{cases}0, & t=0,2,\cdots,N_{j-1}-2;\\ W_{j,\frac{t-1}{2}}, & t=1,3,\cdots,N_{j-1}-1.\end{cases}$$

(我们用类似的方式定义 $V^{\uparrow}_{j,t}$)．用式(81a)并反复应用(95a)得到

$$\begin{aligned}\mathbf{X}&=\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1+\mathcal{A}_1^{\mathrm{T}}\mathcal{B}_2^{\mathrm{T}}\mathbf{W}_2+\mathcal{A}_1^{\mathrm{T}}\mathcal{A}_2^{\mathrm{T}}\mathcal{B}_3^{\mathrm{T}}\mathbf{W}_3+\cdots\\&\quad+\mathcal{A}_1^{\mathrm{T}}\cdots\mathcal{A}_{j-1}^{\mathrm{T}}\mathcal{B}_j^{\mathrm{T}}\mathbf{W}_j+\mathcal{A}_1^{\mathrm{T}}\cdots\mathcal{A}_{j-1}^{\mathrm{T}}\mathcal{A}_j^{\mathrm{T}}\mathbf{V}_j.\end{aligned}$$

细节$\mathcal{D}_j$ 和光滑$\mathcal{S}_j$ 是上式的最后两项，即

$$\begin{aligned}\mathcal{D}_j&=\mathcal{A}_1^{\mathrm{T}}\cdots\mathcal{A}_{j-1}^{\mathrm{T}}\mathcal{B}_j^{\mathrm{T}}\mathbf{W}_j=\mathcal{W}_j^{\mathrm{T}}\mathbf{W}_j\\ \mathcal{S}_j&=\mathcal{A}_1^{\mathrm{T}}\cdots\mathcal{A}_{j-1}^{\mathrm{T}}\mathcal{A}_j^{\mathrm{T}}\mathbf{V}_j=\mathcal{V}_j^{\mathrm{T}}\mathbf{V}_j.\end{aligned} \tag{95c}$$

$\mathcal{P}_j$ 的规范正交性告诉我们

$$\|\mathbf{V}_{j-1}\|^2=\|\mathbf{W}_j\|^2+\|\mathbf{V}_j\|^2.$$

因此 $\mathbf{V}_0=\mathbf{X}$，反复应用上式得到

$$\|\mathbf{X}\|^2=\sum_{k=1}^{j}\|\mathbf{W}_k\|^2+\|\mathbf{V}_j\|^2. \tag{95d}$$

练习[95] 利用上面的式子和式(95c)，证明我们还有

$$\|\mathbf{X}\|^2=\sum_{k=1}^{j}\|\mathcal{D}_k\|^2+\|\mathcal{S}_j\|^2.$$

◁

和$\{X_t\}$的第 j 层的小波系数$\{W_{j,t}\}$相关的等价滤波器$\{h_{j,l}\}$是由下面 j 个滤波器的卷积得到的：

$$\begin{aligned}&\text{滤波器 }1: g_0,g_1,\cdots,g_{L-2},g_{L-1};\\&\text{滤波器 }2: g_0,0,g_1,0,\cdots,g_{L-2},0,g_{L-1};\\&\text{滤波器 }3: g_0,0,0,0,g_1,0,0,0,\cdots,g_{L-2},0,0,0,g_{L-1};\\&\qquad\vdots\\&\text{滤波器 }j-1: g_0,\underbrace{0,\cdots,0}_{2^{j-2}-1\text{个}0},g_1,\underbrace{0,\cdots,0}_{2^{j-2}-1\text{个}0},\cdots,g_{L-2},\underbrace{0,\cdots,0}_{2^{j-2}-1\text{个}0},g_{L-1};\\&\text{滤波器 }j: h_0,\underbrace{0,\cdots,0}_{2^{j-1}-1\text{个}0},h_1,\underbrace{0,\cdots,0}_{2^{j-1}-1\text{个}0},\cdots,h_{L-2},\underbrace{0,\cdots,0}_{2^{j-1}-1\text{个}0},h_{L-1}.\end{aligned} \tag{95e}$$

注意对于 $k=2,\cdots,j-1$，第 k 个滤波器是对前一个滤波器元素之间插入零得到的．

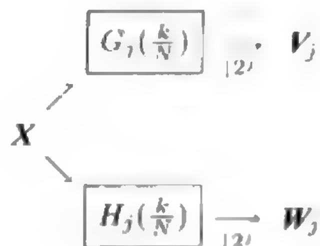


图 96 从 $\mathbf{X}$ 直接计算 $\mathbf{W}_j$ 和 $\mathbf{V}_j$ 的流程图. 长为 N 的时间序列用第 j 层小波滤波器$\{h_{j,l}\}$(传递函数为 $H_j(\cdot)$)进行循环滤波，得到一个输出，然后保留所有的指数可以被 2^j 整除的系数，形成长为 $N/2^j$ 的向量 $\mathbf{W}_j$，即第 j 层小波系数("$\downarrow 2^j$ 表示每 2^j 个系数保留一个"的下采样)；类似地，长为 $N/2^j$ 的向量 $\mathbf{V}_j$ 包含了第 j 层尺度系数，它是对 $\mathbf{X}$ 用传递函数为 $G_j(\cdot)$的尺度滤波器进行循环滤波再进行因子为 2^j 的下采样得到的

练习[96] 证明滤波器$\{h_{j,l}\}$的宽度为

$$L_j \equiv (2^j - 1)(L - 1) + 1, \tag{96a}$$

并且其传递函数为

$$H_j(f) \equiv H(2^{j-1} f) \prod_{l=0}^{j-2} G(2^l f). \tag{96b}$$

◁

滤波器$\{h_{j,l}\}$的标称通带为 $1/2^{j+1} \leqslant |f| \leqslant 1/2^j$，且有

$$2^{j/2} \widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} h_{j,l} X_{t-l \bmod N}, \quad t = 0, \cdots, N-1, \tag{96c}$$

我们有

$$W_{j,t} = 2^{j/2} \widetilde{W}_{j,2^j(t+1)-1} = \sum_{l=0}^{L_j-1} h_{j,l} X_{2^j(t+1)-1-l \bmod N}, \tag{96d}$$

$t=0, 1, \cdots, N_j-1$. 因为我们还有 $\mathbf{W}_j = \mathcal{W}_j \mathbf{X}$，很明显 $\mathcal{W}_j$ 的行包含$\{h^\circ_{j,l}\}$，由定义知它是$\{h_{j,l}\}$以 N 为周期进行周期化的结果.

同样地，和$\{X_t\}$的第 j 层的尺度系数$\{V_{j,t}\}$相关的等价滤波器$\{g_{j,l}\}$是由下面 j 个滤波器的卷积得到的：

$$\begin{aligned}
&\text{滤波器 } 1: g_0, g_1, \cdots, g_{L-2}, g_{L-1};\\
&\text{滤波器 } 2: g_0, 0, g_1, 0, \cdots, g_{L-2}, 0, g_{L-1};\\
&\text{滤波器 } 3: g_0, 0,0,0, g_1, 0,0,0, \cdots, g_{L-2}, 0,0,0, g_{L-1};\\
&\qquad\vdots\\
&\text{滤波器 } j-1: g_0, \underbrace{0,\cdots,0}_{2^{j-2}-1\text{个}0}, g_1, \underbrace{0,\cdots,0}_{2^{j-2}-1\text{个}0}, \cdots, g_{L-2}, \underbrace{0,\cdots,0}_{2^{j-2}-1\text{个}0}, g_{L-1};\\
&\text{滤波器 } j: g_0, \underbrace{0,\cdots,0}_{2^{j-1}-1\text{个}0}, g_1, \underbrace{0,\cdots,0}_{2^{j-1}-1\text{个}0}, \cdots, g_{L-2}, \underbrace{0,\cdots,0}_{2^{j-1}-1\text{个}0}, g_{L-1}.
\end{aligned} \tag{96e}$$

这个等价滤波器$\{h_{j,l}\}$的宽度为 L_j，其传递函数为

$$G_j(f) \equiv \prod_{l=0}^{j-1} G(2^l f). \tag{97}$$

滤波器$\{g_{j,l}\}$近似于一个通带为$0 \leqslant |f| \leqslant 1/2^{j+1}$的低通滤波器，由

$$2^{j/2}\widetilde{V}_{j,t} \equiv \sum_{l=0}^{L_j-1} g_{j,l} X_{t-l \bmod N}, \quad t = 0, \cdots, N-1$$

我们有

$$V_{j,t} = 2^{j/2}\widetilde{V}_{j,2^j(t+1)-1} = \sum_{l=0}^{L_j-1} g_{j,l} X_{2^j(t+1)-1-l \bmod N},$$

$t=0$，1，…，N_j-1. 因为我们还有$\mathbf{V}_j = \mathcal{V}_j \mathbf{X}$，很明显$\mathcal{V}_j$的行包含$\{g^{\circ}_{j,l}\}$，由定义知它是$\{g_{j,l}\}$以$N$为周期进行周期化的结果.

我们现在来看看随着j的增大(除1外)高层的尺度滤波器$\{g_{j,l}\}$和小波滤波器$\{h_{j,l}\}$是什么样子的. 图98a给出了D(4)的$g_{j,l}$(左面)和$h_{j,l}$(右面)的图形，尺度标号分别为$j=1$，2，…，7，滤波器长度为$l=0$，…，L_j-1(图中单位脉冲相应的各点用直线连了起来). 我们看到当j不断增大时，$\{g_{j,l}\}$和$\{h_{j,l}\}$收敛于一个特殊的曲线($\{h_{j,l}\}$的极限曲线看起来有点像鲨鱼的鳍). 我们将在11.4和11.6节详细解释这一点. 图98b的图形是Daubechies小波滤波器的长为8的最接近对称(least asymmetric)的小波(参见4.8节). 注意当j增大时，$\{g_{j,l}\}$收敛于一个类似高斯函数的曲线，而$\{h_{j,l}\}$则收敛于类似图3右面的墨西哥帽小波$\psi^{(\mathrm{Mh})}(\cdot)$的曲线. 图99给出了这些滤波器的平方增益函数$\mathcal{H}_j(\cdot)(j=1, \cdots, 4)$和$\mathcal{G}_4(\cdot)$的图形，其中的竖线表示其标称通带范围.

回忆假设$N=2^J$，我们发现塔式算法的第J次迭代得到的输出向量$\mathbf{W}_J=[W_{J,0}]$和$\mathbf{V}_J=[V_{J,0}]$都是一维向量. 于是塔式算法到这一步就结束了，$W_{J,0}$和$V_{J,0}$构成了$\mathbf{W}=\mathcal{W}\mathbf{X}$的最后两个离散小波变换系数$W_{N-2}$和$W_{N-1}$. 系数$W_{N-2}$和$\{X_t\}$在尺度$\tau_J=2^{J-1}=N/2$上的变化有关. 下面这个练习说明我们总是会得到$W_{N-1}=\overline{X}\sqrt{N}$，其中$\overline{X}$是序列$\{X_t\}$的样本均值.

练习[97] 证明离散小波变换矩阵$\mathcal{W}$最后一行$\mathcal{V}_J$的元素都等于$1/\sqrt{N}$. 用这个结果给出$\mathcal{S}_J$的显式表示式. 提示：考虑$\mathcal{V}_J$中各元素和$1/\sqrt{N}$的差的平方和，并且利用练习[76b]的结论$\sum_l g_l = \sqrt{2}$. ◁

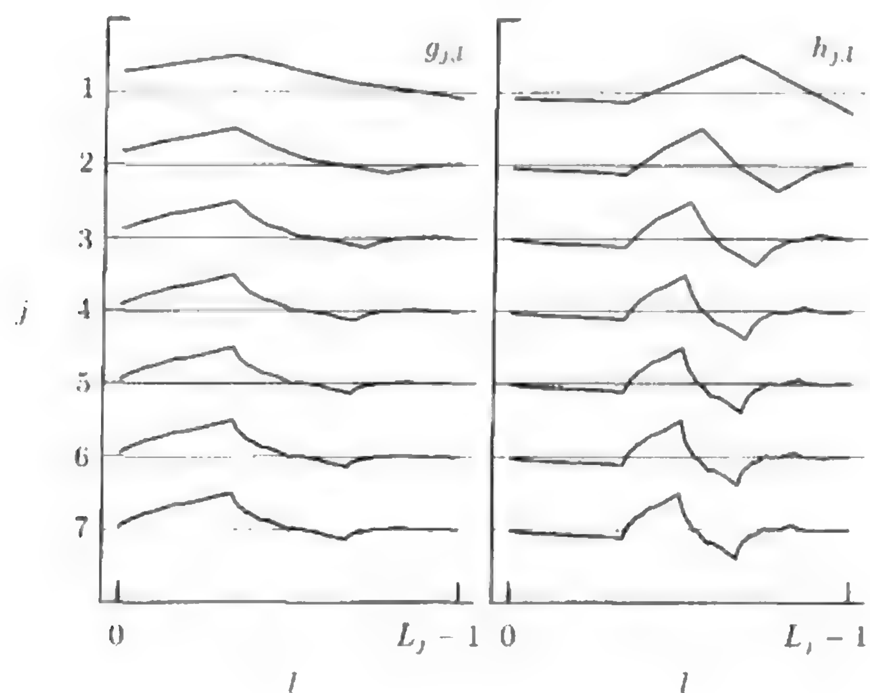


图 98a D(4)尺度滤波器 $\{g_{j,l}\}$和小波滤波器 $\{h_{j,l}\}$，$j=1$，2，…，7(这里脉冲响应序列各个值用直线连了起来)

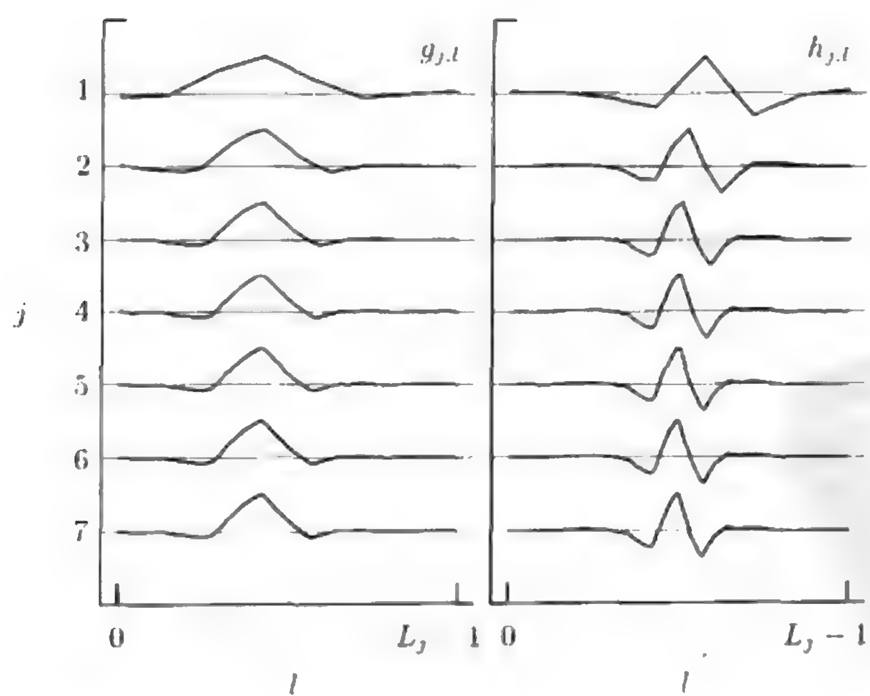


图 98b LA(8)尺度滤波器 $\{g_{j,l}\}$和小波滤波器 $\{h_{j,l}\}$，$j=1$，2，…，7(这里脉冲响应序列各个值用直线连了起来)．这些滤波器将在 4.8 节中定义

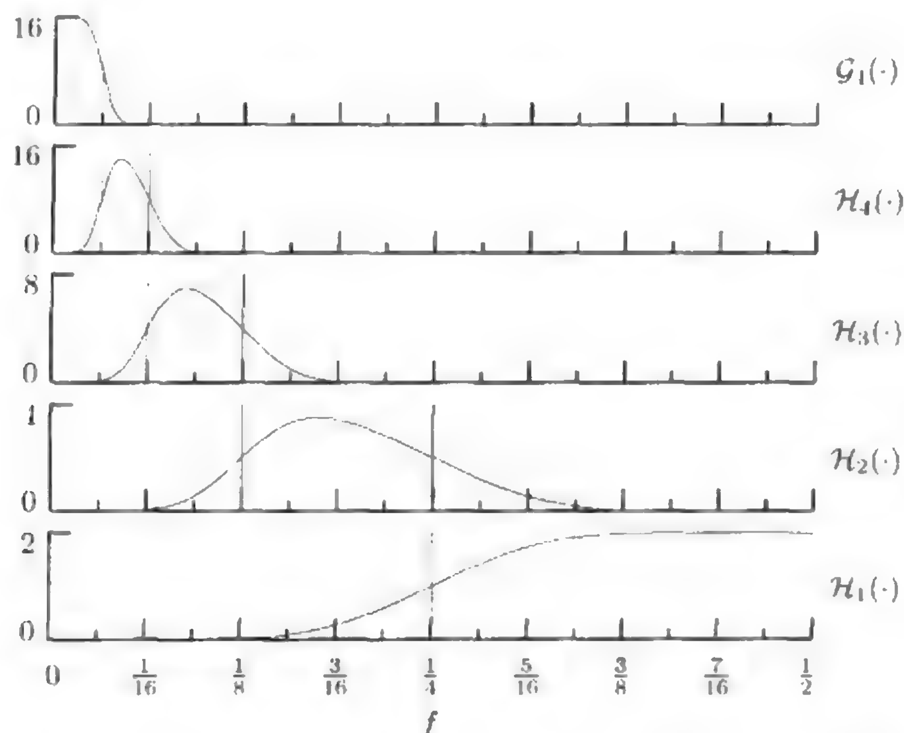


图 99 LA(8)滤波器$\{h_{j,l}\}$，$j=1$，…，4 和$\{g_{4,l}\}$的平方增益函数(这些滤波器将在 4.8 节中定义)

在最后一步中，我们还得到了离散小波变换矩阵的各行表示，即

$$\mathcal{W}=\begin{bmatrix}\mathcal{W}_1\\ \mathcal{W}_2\\ \vdots\\ \mathcal{W}_j\\ \vdots\\ \mathcal{W}_J\\ \mathcal{V}_J\end{bmatrix}=\begin{bmatrix}\mathcal{B}_1\\ \mathcal{B}_2\mathcal{A}_1\\ \vdots\\ \mathcal{B}_j\mathcal{A}_{j-1}\cdots\mathcal{A}_1\\ \vdots\\ \mathcal{B}_J\mathcal{A}_{J-1}\cdots\mathcal{A}_1\\ \mathcal{A}_J\mathcal{A}_{J-1}\cdots\mathcal{A}_1\end{bmatrix}.$$

4.6 节的关键结论与定义

我们前面假设了样本的元素个数为 $N=2^J$，J 为某个整数. 对于 $j=1$，…，J 有 $N_j\equiv N/2^j$，塔式算法的第 j 步将 $N_{j-1}\times N$ 阶矩阵$\mathcal{V}_{j-1}$中的长为 N_{j-1}的行向量分解成两组 N_j 向量，即$\mathcal{W}_j=\mathcal{B}_j\mathcal{V}_{j-1}$和$\mathcal{V}_j=\mathcal{A}_j\mathcal{V}_{j-1}$，其中$\mathcal{B}_j$ 和$\mathcal{A}_j$ 是 $N_j\times N_{j-1}$阶矩阵，分别包含了尺度滤波器$\{h_l\}$和小波滤波器$\{g_l\}$关于周期 N_{j-1}周期化后的序列(这里我们定义$\mathcal{V}_0=I_N$). $\mathcal{A}_j$ 和$\mathcal{B}_j$ 的行向量形成了 N_{j-1}个规范正交向量组. $V_{0,t}\equiv X_t$ 第 j 步变换$\{V_{j-1,t}:t=0,\cdots,N_{j-1}-1\}$为两个新序列，即第 j 层小波系数$\{W_{j,t}:t=0,\cdots,N_j-1\}$和第 j 层尺度系数$\{V_{j,t}:t=0,\cdots,N_j-1\}$. 第 j 层小波系数$\{W_{j,t}\}$与下述结论有关：

[1]$\mathcal{W}_j=\mathcal{B}_j\mathcal{A}_{j-1}\cdots\mathcal{A}_1$，$N_j\times N$ 阶矩阵满足$\mathcal{W}_j\mathcal{W}_j^{\mathrm{T}}=I_{N_j}$，组成了离散小波变换矩阵$\mathcal{W}$从第 $\sum_{k=1}^{j-1}N_k$ 行到第$\sum_{k=1}^{j}N_k-1$行向量($j=1$ 的时候算第 0 行)，$\mathcal{W}_j$ 的各

行包含了$\{h_{j,l}\}$的以 N 为周期的循环平移，即

$$\{h_{j,l}^{\circ}\} \leftrightarrow \left\{H\left(\frac{2^{j-1}k}{N}\right)G\left(\frac{2^{j-2}k}{N}\right)\cdots G\left(\frac{k}{N}\right)\right\}$$

(序列$\{h_{j,l}^{\circ}\}$在$\mathcal{W}_j$ 中是以相反顺序出现的，相邻的两行差一个跨度为 2^j 的平移).

[2]长为 N_j 的向量 $\mathbf{W}_j=\mathcal{W}_j\mathbf{X}$ 的元素.

[3]第 j 层细节$\mathcal{D}_j=\mathcal{W}_j^{\mathrm{T}}\mathbf{W}_j=\mathcal{W}_j^{\mathrm{T}}\ \mathcal{W}_j\mathbf{X}$.

[4]在尺度 $\tau_j=2^{j-1}$ 上的平均变化.

[5]对尺度系数$\{V_{j-1,t}\}$在高频范围为[1/4，1/2]内的部分下 2 采样的结果.

[6]序列$\{X_t\}$的离散傅里叶变换在频率范围为$[1/2^{j+1},\ 1/2^j]$内的部分，是关于频率的一个反序.

相应地，第 j 层尺度系数$\{V_{j,t}\}$与下述结论有关：

[1]$\mathcal{V}_j=\mathcal{A}_j\mathcal{A}_{j-1}\cdots\ \mathcal{A}_1$ 是 $N_j\times N$ 阶矩阵，同时满足$\mathcal{V}_j\mathcal{V}_j^{\mathrm{T}}=I_{N_j}$ 和$\mathcal{W}_j\mathcal{V}_j^{\mathrm{T}}=\mathcal{V}_j\mathcal{W}_j^{\mathrm{T}}=0_{N_j}$，其各行和离散小波变换矩阵$\mathcal{W}$的后 N_j 行张成 $\mathbf{R}^N$ 中同一个子空间. $\mathcal{V}_j$ 的各行是$\{g_{j,l}\}$关于周期 N 循环平移后的结果，即

$$\{g_{j,l}^{\circ}\} \leftrightarrow \left\{G\left(\frac{2^{j-1}k}{N}\right)G\left(\frac{2^{j-2}k}{N}\right)\cdots G\left(\frac{k}{N}\right)\right\}$$

(序列$\{g_{j,l}^{\circ}\}$在$\mathcal{V}_j$ 行中是以反序出现的).

[2]长为 N_j 的向量 $\mathbf{V}_j=\mathcal{V}_j\mathbf{X}$ 的元素.

[3]第 j 层光滑$\mathcal{S}_j=\mathcal{V}_j^{\mathrm{T}}\mathbf{V}_j=\mathcal{V}_j^{\mathrm{T}}\ \mathcal{V}_j\mathbf{X}$.

[4]尺度 $\lambda_j=2^j$ 上的平均.

[5]尺度系数$\{V_{j-1,t}\}$在低频范围[0，1/4]部分的下 2 采样结果.

[6]序列$\{X_t\}$的离散傅里叶变换在频率范围$[0,\ 1/2^{j+1}]$内的部分.

4.6 节的评论与扩展

[1]在这里我们将给出由 Mallat(1989b)引入的计算离散小波变换的塔式算法及其逆运算(重构算法)的伪代码. 令 $V_{0,t}\equiv X_t$，$t=0$，…，$N-1$，其中 $N=2^J$. 令$\{h_n:\ n=0,\ \cdots,\ L-1\}$是一个宽为 L 的小波滤波器，令$\{g_n\}$是由式(75a)定义的相应的尺度滤波器. 如 4.1 节那样，将离散小波变换系数向量 $\mathbf{W}$ 分解成子向量$\mathbf{W}_j$，$j=1$，…，J 和 $\mathbf{V}_J$. 记 $\mathbf{W}_j$ 和 $\mathbf{V}_J$ 的元素是 $W_{j,t}$，$t=0$，…，$N/2^j-1$ 和 $V_{J,t}$，$t=0$，…，$N/2^J-1$. 给定长为 $M\equiv N/2^{j-1}$ 的向量 $\mathbf{V}_{j-1}$，下面的伪代码用来计算长为 $M/2$ 的向量 $\mathbf{W}_j$ 和 $\mathbf{V}_j$(基于式(94a))：

对 $t=0$，… $M/2-1$，进行下面的外循环：

置 $u:=2t+1$.

置 $W_{j,t}:=h_0V_{j-1,u}$并且置 $V_{j,t}:=g_0V_{j-1,u}$.

对于 $n=1$，…，$L-1$，进行下面的内循环：

u 减小 1.

如果 $u<0$，置 $u:=M-1$.

置 $W_{j,t}$等于 $W_{j,t}+h_nV_{j-1,u}$.

置 $V_{j,t}$ 等于 $V_{j,t}+g_n V_{j-1,u}$.

跳出内循环.

跳出外循环.

首先，对于 $j=1$ 利用上面的伪代码(即由 $\boldsymbol{V}_0=\boldsymbol{X}$ 出发)，之后对 $j=2$，…，J，我们得到离散小波变换分量的向量，即 $\boldsymbol{W}_1$，$\boldsymbol{W}_2$，…，$\boldsymbol{W}_J$ 和 $\boldsymbol{V}_J$，以及中间向量 $\boldsymbol{V}_1$，$\boldsymbol{V}_2$，…，$\boldsymbol{V}_{J-1}$.

离散小波逆变换就是通过 $\boldsymbol{W}$ 的子向量 $\boldsymbol{W}_1$，$\boldsymbol{W}_2$，…，$\boldsymbol{W}_J$ 和 $\boldsymbol{V}_J$ 计算 $\boldsymbol{X}$. 给出长为 $M'\equiv N/2^j$ 的向量 $\boldsymbol{W}_j$ 和 $\boldsymbol{V}_j$，下面这段伪代码用来计算长为 $2M'$ 的向量 $\boldsymbol{V}_{j-1}$ (基于式(95b))：

置 $l:=-2$，置 $m:=-1$.

对 $t=0$，… $M'-1$，进行下面的外循环：

置 $l:=l+2, m:=m+2$.

置 $u:=t$，$i:=1$，$k:=0$.

置 $V_{j-1,l}=h_i W_{j,u}+g_i V_{j,u}$.

置 $V_{j-1,m}=h_k W_{j,u}+g_k V_{j,u}$.

如果 $L>2$，那么对于 $n=1$，…，$\frac{L}{2}-1$，进行下面的内循环：

u 增加 1.

如果 $u\geq M'$，置 $u:=0$.

i 增加 2，k 增加 2.

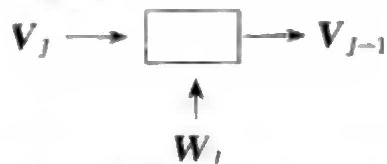
置 $V_{j-1,l}$ 等于 $h_i W_{j,u}+g_i V_{j,u}$.

置 $V_{j-1,m}$ 等于 $h_k W_{j,u}+g_k V_{j,u}$.

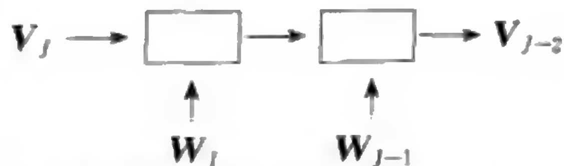
跳出内循环.

跳出外循环.

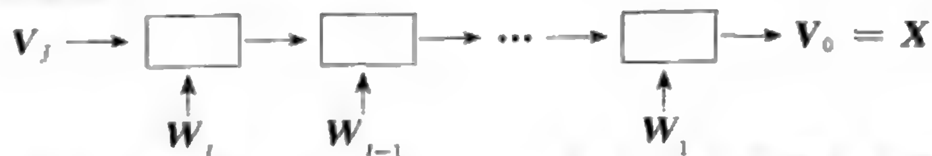
通过对 $\boldsymbol{W}_J$ 和 $\boldsymbol{V}_J$ 用上面的伪代码处理，我们得到 $\boldsymbol{V}_{J-1}$. 将这个过程描述成把 $\boldsymbol{W}_J$ 和 $\boldsymbol{V}_J$ 放进一个盒子里，输出 $\boldsymbol{V}_{J-1}$：



通过对 $\boldsymbol{W}_{J-1}$ 和 $\boldsymbol{V}_{J-1}$ 用上面的伪代码处理，我们得到 $\boldsymbol{V}_{J-2}$. 可以用下面的流程图描述这两步伪代码处理：

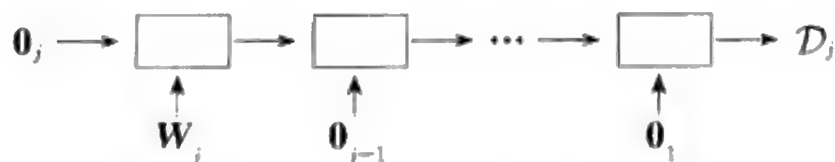


如果按这种方式继续，$J-2$ 次应用上面的伪代码，将会得到 $\boldsymbol{V}_0=\boldsymbol{X}$(注意 $\boldsymbol{V}_{J-3}$，…，$\boldsymbol{V}_1$ 是计算的中间数据)：

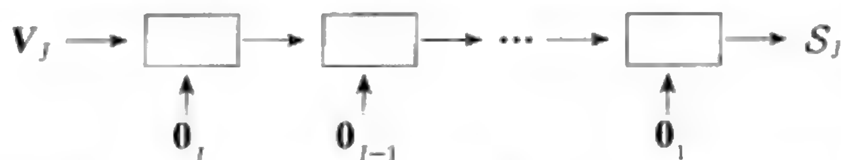


第 j 层细节 $\mathcal{D}_j$ 通过对 $\boldsymbol{0}_1$，…，$\boldsymbol{0}_{j-1}$，$\boldsymbol{W}_j$ 和 $\boldsymbol{0}_j$ 进行离散小波逆变换得到，其中 $\boldsymbol{0}_k$，$k=1$，…，j 是一个有 $N/2^k$ 个零的向量(当 $j=1$ 时，我们解释为 $\boldsymbol{0}_1$，…，$\boldsymbol{0}_{j-1}$，$\boldsymbol{W}_j$ 就是 $\boldsymbol{W}_1$). 这些细节可以依次对 j，$j-1$，…，1 应用逆塔式算法得到，

起始向量是 $\mathbf{W}_j$ 和 $\mathbf{0}_j$（代替了上述伪代码中的 $\mathbf{V}_j$）．在第 j 次迭代中，我们得到 $\mathcal{D}_j$：



类似地，光滑 $\mathcal{S}_J$ 可以通过对 $\mathbf{0}_1$，…，$\mathbf{0}_J$，$\mathbf{V}_J$ 进行离散小波逆变换得到



[2]我们现在讨论一下高层的小波滤波器 $\{h_{j,l}\}$ 和尺度滤波器 $\{g_{j,l}\}$．我们看到，在图 73 中 $\{g_{1,l}\}=\{g_l: l=0, \cdots, L-1\}$ 是一个半带的低通滤波器(是通带为 $0\leqslant|f|\leqslant 1/4$ 的低通滤波器的近似)．从图 91 可以看到 $\{g_{2,l}\}$ 是一个四分之一带的低通滤波器(是通带为 $0\leqslant|f|\leqslant 1/8$ 的低通滤波器的近似)．按照这种方式继续，我们会发现 $\{g_{j,l}\}$ 是一个通带为 $0\leqslant|f|\leqslant 1/2^{j+1}$ 的 2^{-j} 带的低通滤波器．同样会看到，在图 73 中 $\{h_{1,l}\}=\{h_l: l=0, \cdots, L-1\}$ 是一个半带的高通滤波器，或高通滤波器(是通带为 $1/4\leqslant|f|\leqslant 1/2$ 的高通滤波器的近似)．而从图 91 中可以看到 $\{h_{2,l}\}$ 是一个四分之一带的带通滤波器(是通带为 $1/8\leqslant|f|\leqslant 1/4$ 的高通滤波器的近似)．类似地，$\{h_{j,l}\}$ 是一个通带为 $1/2^{j+1}\leqslant|f|\leqslant 1/2^j$ 的 2^{-j} 带的(倍频程)带通滤波器．

有两个等价的方法可以得到 $\{h_{j,l}\}$ 和 $\{g_{j,l}\}$．第一种方法：通过在 $\{h_l\}$ 和 $\{g_l\}$ 元素之间分别插入 $2^{j-1}-1$ 个零得到 $\{h_{j,l}\}$ 和 $\{g_{j,l}\}$．然后按下述方法应用 $2^{-(j-1)}$ 带的低通滤波器 $\{g_{j-1,l}\}$(见式(95e)和(96e))．

练习[102a] 给定 $\{g_{j-1,l}\}$，证明可以通过下式得到 $\{h_{j,l}\}$ 和 $\{g_{j,l}\}$：

$$h_{j,l}=\sum_{k=0}^{L-1} h_k g_{j-1,l-2^{j-1}k}, \quad g_{j,l}=\sum_{k=0}^{L-1} g_k g_{j-1,l-2^{j-1}k}, \tag{102}$$

$l=0, \cdots, L_j-1$(参见 Rioul，1992)．还证明，通过定义 $G_1(f)\equiv G(f)$，对任意的 $j\geqslant 2$，有

$$H_j(f)=H(2^{j-1}f)G_{j-1}(f), \quad G_j(f)=G(2^{j-1}f)G_{j-1}(f). \qquad \triangleleft$$

第二种方法：下一层的小波滤波器和尺度滤波器是通过在 $\{h_{j-1,l}\}$ 或 $\{g_{j-1,l}\}$ 中插入单个零得到，然后应用半带低通滤波器 $\{g_l\}$．

练习[102b] 证明

$$h_{j,l}=\sum_{k=0}^{L_{j-1}-1} g_{l-2k} h_{j-1,k}, \quad g_{j,l}=\sum_{k=0}^{L_{j-1}-1} g_{l-2k} g_{j-1,k},$$

$l=0, \cdots, L_j-1$. $\triangleleft$

注意$\{h_{j,l}\}$的建立不涉及$\{g_l\}$和$\{h_{j-1,l}\}$的卷积(将上式中的 $2k$ 用 k 代替即可). 将$\{h_{j-1,l}\}$映射为$\{h_{j,l}\}$的过程在工程中有时候被称为是"插值算子"(参见 Rioul and Duhamel, 1992). 这个术语在这里是适用的, 因为我们用低通滤波器$\{g_l\}$对新的滤波器进行插值, 将尺度为 τ_{j-1} 的滤波器$\{h_{j-1,l}\}$"拉长"成尺度为 $\tau_j=2\tau_{j-1}$.

[3]现在证明我们的结论: 尺度滤波器$\{g_{j,l}\}$对应于尺度 $\lambda_j=2^j$. 虽然$\{g_{j,l}\}$是由 $L_j=(2^j-1)(L-1)+1$ 给出, 由图 98a 和图 98b 可以看到, 随着 j 的增大, $g_{j,l}$在 $l=0$ 和 $l=L_j-1$ 处的很多值非常接近于零, 因此用 L_j 来度量$\{g_{j,l}\}$的宽度显得有点不合适. 一个更好的表示序列宽度的度量是基于"自相关"宽度的概念的, 自相关宽度在这里定义如下:

$$\text{width}_a\{g_{j,l}\} \equiv \frac{\left(\sum_{l=-\infty}^{\infty} g_{j,l}\right)^2}{\sum_{l=-\infty}^{\infty} g_{j,l}^2} \tag{103}$$

(详情参见 Bracewell, 1978 或 Percival and Walden, 1993). 考虑前 j 层哈尔尺度滤波器:

$$g_{j,l}=\begin{cases}1/2^{j/2}, & l=0,\cdots,2^j-1;\\ 0, & \text{其他}.\end{cases}$$

这个滤波器的输出值和有 2^j 个输入的样本均值是成比例的, 因此哈尔$\{g_{j,l}\}$很明显对应于尺度 $2^j=\lambda_j$.

练习[103] 证明哈尔$\{g_{j,l}\}$的自相关宽度为 2^j, 直观上这很自然. ◁

对于所有的尺度滤波器, 我们有

$$\sum_{l=-\infty}^{\infty} g_{j,l} = \sum_{l=0}^{L_j-1} g_{j,l} = G_j(0) = \prod_{k=0}^{j-1} G(0) = 2^{j/2},$$

由于 $G(0)=\sum_{l=0}^{L-1} g_l=\sqrt{2}$, 而

$$\sum_{l=0}^{L_j-1} g_{j,l}^2 = 1,$$

因为对任意的 $N\geq L_j$, 滤波器系数$\{g_{j,l}\}$是$\mathcal{V}_j$ 中行的非零元素, 而$\mathcal{V}_j$ 满足规范正交性质$\mathcal{V}_j\mathcal{V}_j^T=I_{N_j}$, 因此可得 $\text{width}_a\{g_{j,l}\}=2^j=\lambda_j$.

进而, 因为小波滤波器$\{h_{j,l}\}$和$\{g_{j,l}\}$的宽度都为 L_j, 两个滤波器正交, 且小波滤波器的系数和为零. 每个小波系数可以看成是广义的局部平均后进行差分的结果, 每个系数的有效长度都是 λ_j 的一半, 即和$\{h_{j,l}\}$的相对应的尺度 $\lambda_j/2=\tau_j$, 这个结果我们前面已经一直在用了.

4.7 部分离散小波变换

对于时间序列$\{X_t: t=0, \cdots, N-1\}$，样本的个数为$N=2^J$(如前面所假设的那样)，塔式算法$J$次迭代后导出了离散小波变换系数$\mathbf{W}=\mathcal{W}\mathbf{X}$. 如果在$J_0<J$步后就停止，就得到了$\mathbf{X}$的$J_0$层的部分离散小波变换系数

$$\begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \\ \vdots \\ \mathbf{W}_j \\ \vdots \\ \mathbf{W}_{J_0} \\ \mathbf{V}_{J_0} \end{bmatrix} = \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{W}_2 \\ \vdots \\ \mathcal{W}_j \\ \vdots \\ \mathcal{W}_{J_0} \\ \mathcal{V}_{J_0} \end{bmatrix} \mathbf{X} = \begin{bmatrix} \mathcal{B}_1 \\ \mathcal{B}_2\mathcal{A}_1 \\ \vdots \\ \mathcal{B}_j\mathcal{A}_{j-1}\cdots\mathcal{A}_1 \\ \vdots \\ \mathcal{B}_{J_0}\mathcal{A}_{J_0-1}\cdots\mathcal{A}_1 \\ \mathcal{A}_{J_0}\mathcal{A}_{J_0-1}\cdots\mathcal{A}_1 \end{bmatrix} \mathbf{X}.$$

这里$\mathbf{W}_j$，$j=1, \cdots, J_0$，是离散小波变换系数向量$\mathbf{W}$的子向量，最后的那个子向量为$N/2^{J_0}$个尺度系数的最后子向量$\mathbf{V}_{J_0}$，代替$\mathbf{W}$的后$N/2^{J_0}$个系数. 所有尺度系数代表在尺度$\lambda_{J_0}=2^{J_0}$上的平均，因此包含了$\mathbf{X}$在大尺度(或低频)上的特征(包括样本平均$\overline{X}$). 作为比较，完全离散小波变换只有一个尺度系数，这就只包含了样本均值(参见练习[97]).

部分离散小波变换在实际中应用比完全离散小波变换更为普遍，我们可以只分解到指定的某一尺度，因为更大尺度上的小波系数我们并不关心. J_0层的部分离散小波变换使得我们可以放宽以前的限制$N=2^J$，J为某一整数，现在N只需是2^{J_0}的整数倍就行了. 采样的数量并不是我们要考虑的最重要的因素，J_0可以根据特定的应用来确定(参见 4.10 和 5.8 节).

对于分解层数$J_0<J$的部分离散小波变换，我们得到如下分解(即多分辨分析)：

$$\mathbf{X}=\sum_{j=1}^{J_0}\mathcal{D}_j+\mathcal{S}_{J_0}, \tag{104a}$$

如前所述，细节$\mathcal{D}_j$表示在尺度$\tau_j=2^{j-1}$上的变化，而$\mathcal{S}_{J_0}$表示在尺度$\lambda_{J_0}=2^{J_0}$上的平均. 能量分解为

$$\|\mathbf{X}\|^2=\sum_{j=1}^{J_0}\|\mathbf{W}_j\|^2+\|\mathbf{V}_{J_0}\|^2=\sum_{j=1}^{J_0}\|\mathcal{D}_j\|^2+\|\mathcal{S}_{J_0}\|^2, \tag{104b}$$

由于样本方差为$\hat{\sigma}_X^2=\frac{1}{N}\|\mathbf{X}\|^2-\overline{X}^2$，所以样本方差的分解如下：

$$\hat{\sigma}_X^2=\frac{1}{N}\sum_{j=1}^{J_0}\|\mathbf{W}_j\|^2+\frac{1}{N}\|\mathbf{V}_{J_0}\|^2-\overline{X}^2=\frac{1}{N}\sum_{j=1}^{J_0}\|\mathcal{D}_j\|^2+\frac{1}{N}\|\mathcal{S}_{J_0}\|^2-\overline{X}^2, \tag{104c}$$

其中上式的最后两项可以认为是J_0层光滑$\mathcal{S}_{J_0}$的样本方差(参见练习[4.20]).

4.8 Daubechies 小波滤波器和尺度滤波器：形式和相位

我们在前面几节已经看到，可以根据小波滤波器满足的性质来构造正交离散小波变换矩阵$\mathcal{W}$，即求和为零性质(式(69a))和规范正交性质(式(69b)和(69c)). 对于哈尔和 D(4)小波滤波器，我们从图 73 可以看出他们都是标称通带为[1/4, 1/2]的高通滤波器. 这个事实表明相应的尺度滤波器是低通滤波器，并且第 j 层小波系数相应的频率在通带$[1/2^{j+1}, 1/2^j]$上. 值得注意的一点是，小波滤波器的这个带通特性不是小波滤波器定义中性质的推论，如下述练习所表明的.

练习[105a] 证明滤波器$\{a_0=-1/\sqrt{2}, a_1=0, a_2=0, a_3=1/\sqrt{2}\}$是一个小波滤波器. 计算它的平方增益函数，论证这个小波滤波器不是高通滤波器的一个适当的逼近. ◁

上面这个练习中的小波滤波器除了不是一个高通滤波器外，用这个滤波器进行小波变换得到的小波系数不能像哈尔小波滤波器和 D(4)小波滤波器那样可以将结果解释为先将相邻加权平均后再进行(广义)差分得到. 因此如果想用类似前面的方式来解释得到的小波系数，还需要对小波滤波器加入额外的限制条件.

通过加入一组正则性条件(将在 11.9 节讨论)，Daubechies(1992, 6.1 节)得到了一类很有用的小波滤波器，这类小波滤波器生成的离散小波变换都可以表示成对相邻平均做差分. 这类小波滤波器可以用 Daubechies 尺度滤波器$\{g_l: l=0, \cdots, L-1\}$的平方增益函数来定义：

$$\mathcal{G}^{(D)}(f) \equiv 2\cos^L(\pi f)\sum_{l=0}^{\frac{L}{2}-1}\binom{\frac{L}{2}-1+l}{l}\sin^{2l}(\pi f), \tag{105a}$$

其中 L 是一个正的偶数，

$$\binom{a}{b} \equiv \frac{a!}{b!(a-b)!}, \quad \sin^0(\pi f)=1 \text{ 对所有的 } f \text{ 成立}.$$

用关系$\mathcal{H}^{(D)}(f)=\mathcal{G}^{(D)}\left(f+\frac{1}{2}\right)$，我们可以看到 Daubechies 小波滤波器的平方增益函数满足

$$\mathcal{H}^{(D)}(f) \equiv 2\sin^L(\pi f)\sum_{l=0}^{\frac{L}{2}-1}\binom{\frac{L}{2}-1+l}{l}\cos^{2l}(\pi f). \tag{105b}$$

就像下面将要指出的那样，我们可以将$\mathcal{H}^{(D)}(\cdot)$作为小波滤波器的平方增益函数.

练习[105b] 证明对于差分滤波器$\{a_0=1, a_1=-1\}$，其平方增益函数为$\mathcal{D}(f)\equiv 4\sin^2(\pi f)$. ◁

利用这个结果，我们可以写成$\mathcal{H}^{(D)}(f)=\mathcal{D}^{\frac{L}{2}}(f)\mathcal{A}_L(f)$，其中

$$\mathcal{A}_L(f)\equiv\frac{1}{2^{L-1}}\sum_{l=0}^{\frac{L}{2}-1}\binom{\frac{L}{2}-1+l}{l}\cos^{2l}(\pi f).$$

图 107 绘出了对于$L=2$，4，…，14 的$\mathcal{D}^{\frac{L}{2}}(\cdot)$、$\mathcal{A}_L(\cdot)$和$\mathcal{H}^{(D)}(\cdot)$的图形．对于$L\geq 4$，我们看到滤波器$\mathcal{A}_L(\cdot)$近似于一个低通滤波器，因此它的作用就是光滑或加权平均(对于$L=2$，即哈尔小波的情形，这个滤波器是全通滤波器)．

练习[106] 证明式(105b)的平方增益函数$\mathcal{H}^{(D)}(\cdot)$对所有的正偶数L满足$\mathcal{H}^{(D)}(f)+\mathcal{H}^{(D)}\left(f+\frac{1}{2}\right)=2$(回忆这个条件保证了任何具有平方增益函数$\mathcal{H}^{(D)}(\cdot)$的小波滤波器$\{h_l\}$具有单位能量并且与它的偶平移正交)．提示：可以通过数学归纳法证明． ◁

随着L的增大，$\mathcal{G}^{(D)}(\cdot)$收敛于理想低通滤波器的平方增益函数(Lai，1995)，并且满足平方增益函数为$\mathcal{G}^{(D)}(\cdot)$的不同的实值滤波器$\{g_l:\ l=0,\cdots,L-1\}$也越来越多(将在 11.10 节详细讨论这样的滤波器的个数)．这些不同的$\{g_l\}$的传递函数$G(\cdot)$一定也是不一样的，不过差别仅仅是他们的相位函数不同，即$\theta^{(G)}(\cdot)$不同，用极坐标表示为

$$G(f)=[\mathcal{G}^{(D)}(f)]^{1/2}e^{i\theta^{(G)}(f)}.$$

给定$\mathcal{G}^{(D)}(\cdot)$，我们可以通过谱分解的方式得到所有可能的$\{g_l\}$(将在 11.10 节详细介绍)，不同的分解将得到不同的相位函数$\theta(\cdot)$．那么自然会有这样的问题：在这么多分解方式中我们采用哪种分解好呢？Daubechies 最初用的分解方式是选择极小相位，得到了工程中称为是最小延迟的滤波器(Oppenheim and Schafer，1989，5.6 节)．如果用$\{g_l^{(ep)}\}$记极小相位滤波器，设$\{g_l\}$是另一个种分解得到的滤波器，则有

$$\sum_{l=0}^{m}g_l^2\leqslant\sum_{l=0}^{m}[g_l^{(ep)}]^2,\quad m=0,\cdots,L-1$$

(Oppenheim and Schafer，1989，式(5.118))．注意$m=L-1$时等式成立，因为所有的滤波器都有单位能量．对于$m<L-1$，类似上面的求和定义了一个部分能量序列，这个序列可以得到最小延迟滤波器．

我们称 Daubechies 的长为L的极小相位滤波器为 D(L)尺度滤波器，$L=2$，4，…(这些滤波器被 Bruce and Gao，1996a 称为是“daublets”)．D(2)尺度函数就是哈尔尺度函数，对于这种情形，$\{g_0^{(ep)}=1/\sqrt{2}$且$g_1^{(ep)}=1/\sqrt{2}\}$；D(4)尺度滤波器系数在式(75d)中已经给出了，D(6)和 D(8)滤波器系数在表 109 中给出(这个表的数据来自 Daubechies，1992，p.195 表 6.1)．Daubechies 的$L=2$，4，…，20 的极小相位滤波器在图 108a 中给出了，而相应的小波滤波器(由式(75b)得到)在图 108b 中给出．

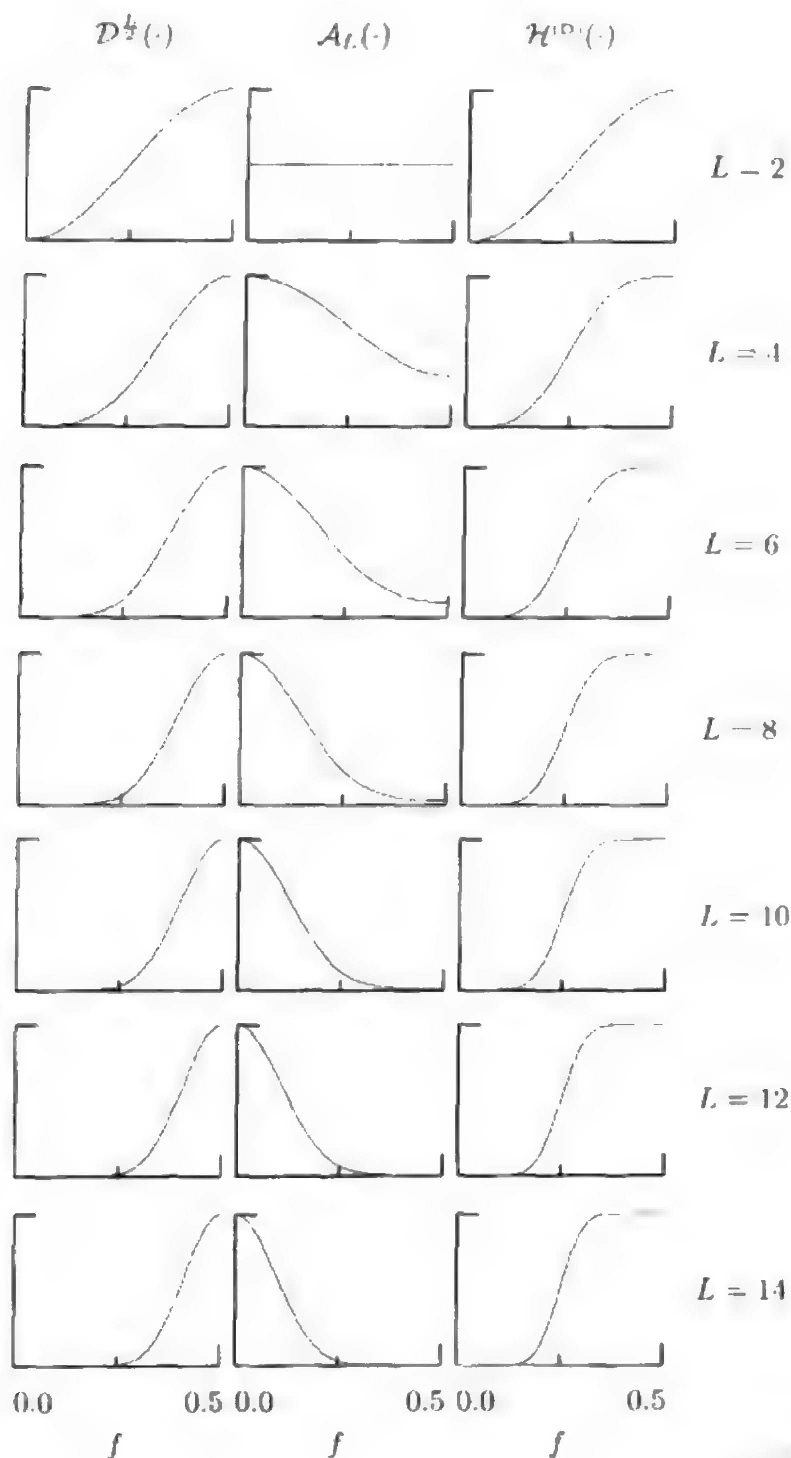


图 107 $L=2$, 4, …, 14 的 Daubechies 小波滤波器的平方增益函数 $\mathcal{H}^{(D)}(\cdot)$（右列的图）. 每个 $\mathcal{H}^{(D)}(\cdot)$ 是另外两个平方增益函数的乘积，即 $\mathcal{D}^{\frac{L}{2}}(\cdot)$（左列图）和 $\mathcal{A}_L(\cdot)$（中图）的乘积. 第一个滤波器是 $\frac{L}{2}$ 阶差分滤波器，当 $L\geqslant4$ 时第二个滤波器是一个加权平均滤波器（即低通滤波器）

第二种分解 $\mathcal{G}^{(D)}(\cdot)$ 的选择得到了最接近对称（LA）的一族尺度滤波器，我们记为 $\{g_l^{(la)}\}$（这些还被称为"symmlets"，见 Bruce and Gao 1996a，或 Härdle 等，1998）. 为了理解这种分解的原理，我们需要引入零相位滤波器和线性相位滤波

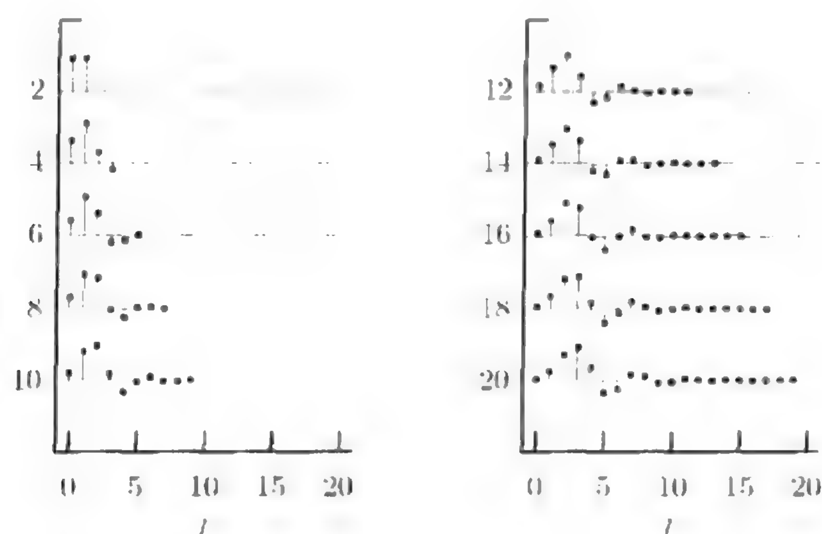


图 108a Daubechies 极小相位尺度滤波器 $\{g_l^{(\mathrm{ep})}\}$，$L=2, 4, \cdots, 20$(基于 Daubechies 的值，1992，p. 195，表 6.1)

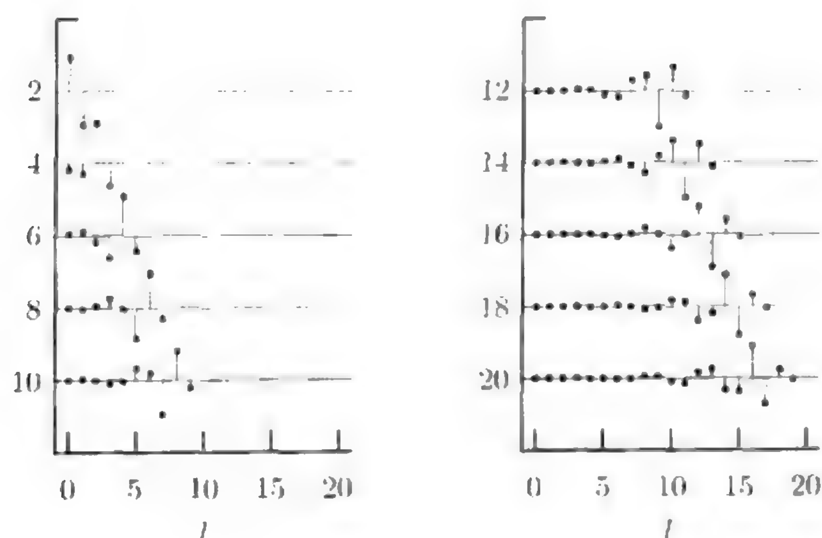


图 108b 相应于图 108a 的极小相位尺度滤波器的 Daubechies 小波滤波器

器的概念. 考虑滤波器$\{u_l\}$的传递函数为

$$U(f)\equiv\sum_{l=-\infty}^{\infty}u_l e^{-i2\pi fl}.$$

令$\{u_l^{\circ}: l=0, \cdots, N-1\}$是$\{u_l\}$以 N 为周期进行周期化的结果，并且令$\{U_k^{\circ}\}$是它的离散傅里叶变换：

$$U_k^{\circ}\equiv\sum_{l=0}^{N-1}u_l^{\circ}e^{-i2\pi kl/N}=U(f_k),\quad f_k\equiv\frac{k}{N}.$$

表 109

l	对 D(6)的 g_l	对 C(6)的 g_l	对 D(8)的 g_l
0	0.332 670 552 950 082 7	−0.015 655 728 528 984 8	0.230 377 813 307 443 1
1	0.806 891 509 311 092 8	−0.072 732 621 341 051 1	0.714 846 570 548 405 8
2	0.459 877 502 118 491 5	0.384 864 856 538 113 4	0.630 880 767 935 878 8
3	−0.135 011 020 010 254 6	0.852 572 041 642 390 0	−0.027 983 769 416 683 4
4	−0.085 441 273 882 026 7	0.337 897 670 951 159 0	−0.187 034 811 717 913 2
5	0.035 226 291 885 709 6	−0.072 732 275 741 188 9	0.030 841 381 835 366 1
6			0.032 883 011 666 677 8
7			−0.010 597 401 785 002 1

l	对 LA(8)的 g_l	对 LA(12)的 g_l	对 LA(16)的 g_l
0	−0.075 765 714 789 340 7	0.015 404 109 327 337 7	−0.003 382 415 951 359 4
1	−0.029 635 527 645 954 1	0.003 490 712 084 330 4	−0.000 542 132 331 635 5
2	0.497 618 667 632 457 8	−0.117 990 111 148 410 5	0.031 695 087 810 345 2
3	0.803 738 751 805 216 3	−0.048 311 742 585 998 1	0.007 607 487 325 284 8
4	0.297 857 795 605 542 2	0.491 055 941 927 639 6	−0.143 294 238 351 054 2
5	−0.099 219 543 576 935 4	0.787 641 141 028 794 1	−0.061 273 359 067 908 8
6	−0.012 603 967 262 261 2	0.337 929 421 728 240 1	0.481 359 651 259 201 2
7	0.032 223 100 604 071 3	−0.072 637 522 786 600 0	0.777 185 751 699 747 8
8		−0.021 060 292 512 695 4	0.364 441 894 835 956 4
9		0.044 724 901 770 748 2	−0.051 945 838 107 875 1
10		0.001 767 711 864 398 3	−0.027 219 029 916 813 7
11		−0.007 800 708 324 765 0	0.049 137 179 673 476 8
12			0.003 808 752 014 060 1
13			−0.014 952 258 336 792 6
14			−0.000 302 920 514 551 6
15			0.001 889 950 332 900 7

注：一些 Daubechies 尺度滤波器的系数和 coiflet 尺度滤波器(将在 4.9 节讨论)的系数. 这个表中的系数来自 Daubechies(1992, 1993). 这些系数(还有其他的滤波器 $\{g_l\}$)在本书的网站上可以找到.

考虑时间序列 $\{X_t:\ t=0,\cdots,N-1\}$, 其离散傅里叶变换为

$$\mathcal{X}_k \equiv \sum_{t=0}^{N-1} X_t \mathrm{e}^{-\mathrm{i}2\pi kt/N}, \quad \text{因此}\ X_t = \frac{1}{N}\sum_{k=0}^{N-1} \mathcal{X}_k \mathrm{e}^{\mathrm{i}2\pi kt/N}.$$

令

$$Y_t \equiv \sum_{l=0}^{N-1} u_l^{\circ} X_{t-l \bmod N}, \quad t=0,\cdots,N-1$$

是用 $\{u_l^{\circ}\}$ 对 $\{X_t\}$ 进行循环滤波的结果. 于是 $\{Y_t\}$ 的离散傅里叶变换为 $\{U_k^{\circ}\mathcal{X}_k\}$, 离散傅里叶逆变换告诉我们

$$Y_t = \frac{1}{N}\sum_{k=0}^{N-1} U_k^{\circ}\mathcal{X}_k \mathrm{e}^{\mathrm{i}2\pi kt/N}.$$

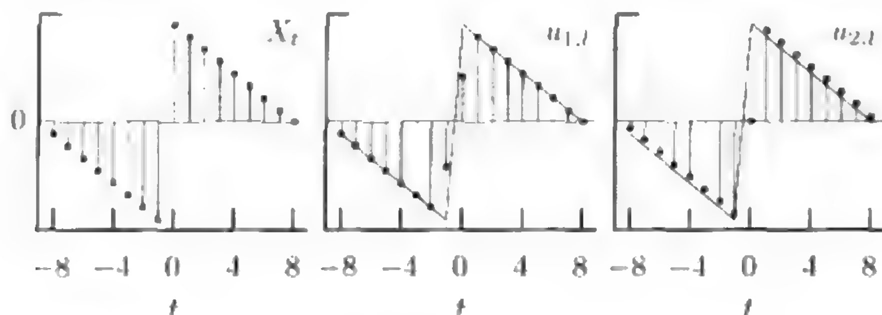


图 110 用零相位滤波器滤波的例子. 左图给出了一个时间序列$\{X_t\}$，它在$t=-1$和$t=0$处不连续. 中图和右图的实心点表示对这个序列分别用滤波器$\{u_{1,l}\}$(式(110a))和$\{u_{2,l}\}$(式(110b))进行滤波的结果. 只有第一个滤波器是具有零相位的. 这两幅图中的细线表示原始的时间序列. 注意用零相位滤波器$\{u_{1,l}\}$不连续点发散开了——但在位置上没有平移. 但是用$\{u_{2,l}\}$滤波结果使得不连续点发散了且在位置上有一个平移

让我们把传递函数$U(\cdot)$写成$U(f)=|U(f)|e^{i\theta(f)}$，其中$\theta(\cdot)$是相位函数. 如果$\theta(f_k)=0$对所有的$k$成立，那么$U(f_k)=|U(f_k)|$，于是$U_k^\circ=|U_k^\circ|$，所以我们有

$$Y_t=\frac{1}{N}\sum_{k=0}^{N-1}|U_k^\circ|\mathcal{X}_k e^{i2\pi kt/N},$$

因此用$\{u_l^\circ\}$对$\{X_t\}$进行滤波得到一个序列，它的重构和序列$\{X_t\}$只在幅度上有差别，相位是一样的. $\theta(f_k)=0$的实际含义是事件$\{Y_t\}$和事件$\{X_t\}$可以"对齐". 这在物理上解释$\{Y_t\}$的意义的时候是非常有用的. 一个滤波器的相位函数对所有的f满足$\theta(f)=0$时称为是零相位滤波器.

下面是一个零相位滤波器的例子：

$$u_{1,l}=\begin{cases}1/2, & l=0;\\ 1/4, & l=\pm 1;\\ 0, & \text{其他}.\end{cases} \tag{110a}$$

从练习[2.3]，我们有$U_1(f)=\cos^2(\pi f)$是实值非负的. 并且对所有的f有$\theta_1(f)=0$. 作为对比，考虑滤波器

$$u_{2,l}=\begin{cases}1/2, & l=0,1;\\ 0, & \text{其他}.\end{cases} \tag{110b}$$

其传递函数为

$$U_2(f)=\frac{1}{2}+\frac{e^{-i2\pi f}}{2}=e^{-i\pi f}\cos(\pi f)=|U_2(f)|e^{-i\pi f}.$$

由于$\theta_2(f)=-\pi f$，所以这个滤波器不具有零相位. 这两个滤波器滤波时间序列的例子见图 110.

为了引入线性相位滤波器，现在考虑实现间隔为 ν 单位的循环推进滤波器输出$\{Y_t\}$，形如

$$Y_t^{(\nu)} \equiv Y_{t+\nu \bmod N}, \quad t = 0, \cdots, N-1,$$

其中 ν 是一个整数且满足 $1 \leqslant |\nu| \leqslant N-1$(用向量的记号$\{Y_t^{(\nu)}\}$形成了 $\mathcal{T}^{-\nu}\boldsymbol{Y}$ 的元素，参见关于式(52a)的讨论)．举个例子，如果 $\nu=2$ 且 $N \geqslant 11$，那么 $Y_8^{(2)}=Y_{10}$，因此$\{Y_t^{(2)}\}$中的事件比$\{Y_t\}$要早两个事件单位．现在

$$\begin{aligned} Y_t^{(\nu)} = Y_{t+\nu \bmod N} &= \sum_{l=0}^{N-1} u_l^{\circ} X_{t+\nu-l \bmod N} \\ &= \sum_{l=-\nu}^{N-1-\nu} u_{l+\nu}^{\circ} X_{t-l \bmod N} \\ &= \sum_{l=0}^{N-1} u_{l+\nu \bmod N}^{\circ} X_{t-l \bmod N}. \end{aligned}$$

于是将滤波输出提前 ν 个单位相当于用一个系数前移 ν 单位的滤波器．周期滤波器$\{u_{l+\nu \bmod N}^{\circ}: l=0, \cdots, N-1\}$可以通过将滤波器$\{u_l^{(\nu)} \equiv u_{l+\nu}: l=\cdots, -1, 0, 1, \cdots\}$以 N 为周期进行周期化．这样由$\{u_l^{(\nu)}\}$的相位性质就去掉，成了新的滤波器$\{u_{l+\nu \bmod N}^{\circ}\}$．

练习[111] 证明前移滤波器$\{u_l^{(\nu)}\}$的传递函数为

$$U^{(\nu)}(f) \equiv e^{i2\pi f\nu} U(f). \qquad \triangleleft$$

因此如果$\{u_l\}$具有零相位，即 $U(f) = |U(f)|$，那么 $U^{(\nu)}(f) = |U(f)| \exp(i2\pi f\nu)$，于是$\{u_l^{(\nu)}\}$的相位函数为

$$\theta(f) = 2\pi f\nu.$$

对某个实值常数 ν，相位函数满足上式的滤波器称为是线性相位滤波器．如果 ν 是一个整数，那么线性相位滤波器很容易就可以变成零相位滤波器．

作为一个例子，假定

$$u_{3,l} = \begin{cases} 1/2, & l = 1; \\ 1/4, & l = 0 \text{ 或 } 2; \\ 0, & \text{其他}. \end{cases}$$

这个滤波器的传递函数为

$$U_3(f) \equiv \frac{1}{4} + \frac{e^{-i2\pi f}}{2} + \frac{e^{-i4\pi f}}{4} = \cos^2(\pi f) e^{-i2\pi f},$$

显然是线性相位的，且 $\nu=-1$．于是将$\{u_{3,l}\}$前移(即左移)1 个单位就变成了式(110a)中的零相位滤波器$\{u_{1,l}\}$；即 $u_{3,l}^{(1)} = u_{1,l}$．

长为L的Daubechies最接近对称的尺度滤波器如下定义. 对于具有平方增益函数$\mathcal{G}^{(D)}(\cdot)$的所有可能的序列$\{g_l: l=0, \cdots, L-1\}$，给定一个平移$\tilde{\nu}$，计算

$$\rho_{\tilde{\nu}}(\{g_l\}) \equiv \max_{-1/2\leqslant f\leqslant 1/2} |\theta^{(G)}(f) - 2\pi f\tilde{\nu}|, \tag{112a}$$

其中$\theta^{(G)}(\cdot)$是我们要考虑的$\{g_l\}$的相位函数(我们必定有$\rho_{\tilde{\nu}}(\{g_l\})>0$，因为紧支撑尺度滤波器不可能具有严格线性相位，参见Daubechies 1992，定理8.1.4). 对于给定的$\{g_l\}$，令ν是使得上式最小的平移(实际上，由于我们的假设$\{g_l\}$的非零值只可能在$l=0$到$L-1$中，所以必有$-(L-1)\leqslant\nu\leqslant 0$)，因此我们只需要计算$\rho_{\tilde{\nu}}(\{g_l\})$的有限个平移). 最接近对称的滤波器$\{g_l^{(la)}\}$使得$\rho_\nu(\{g_l\})$最小. 换句话说，最接近对称的滤波器的相位函数和线性相位偏差最小. 我们以后称Daubechies的长为L的最接近对称的滤波器为LA(L)尺度滤波器，其中$L=8$，10，…. 对于LA(8)、LA(12)和LA(16)滤波器在表109中给出(表中的数据是从Daubechies 1992，第198页，表6.3计算得到，但是我们的数据和Daubechies给出的不一样，原因将在评论与扩展第一条给出). LA(8)，LA(10)，…，LA(20)尺度滤波器在图113a中给出，相应的小波滤波器在图113b中给出.

由于ν是使得式(112a)最小的$\tilde{\nu}$，我们有

$$\theta^{(G)}(f) \approx 2\pi f\nu. \tag{112b}$$

练习[112] 利用式(76a)证明

$$H(f) = e^{-i2\pi f(L-1)+i\pi} G\left(\frac{1}{2} - f\right).$$

◁

于是对于相应的小波滤波器的相位函数为

$$\theta^{(H)}(f) = -2\pi f(L-1) + \pi + \theta^{(G)}\left(\frac{1}{2} - f\right).$$

由式(112b)推出

$$\theta^{(H)}(f) \approx -2\pi f(L-1+\nu) + \pi(\nu+1), \tag{112c}$$

因此，如果ν是奇数，小波滤波器的相位函数还是近似为线性的(回忆相位函数是模2π定义的). 通过计算表明对所有的最接近对称滤波系数ν确实是奇数，因此所有的最接近对称尺度滤波器和小波滤波器的相位函数可以取近似线性相位：

$$\theta^{(G)}(f) \approx 2\pi f\nu, \quad \theta^{(H)}(f) \approx -2\pi f(L-1+\nu), \tag{112d}$$

其中ν总是奇数. 特别地，通过计算表明下式成立：

$$\nu = \begin{cases} -\dfrac{L}{2}+1, & L = 8,12,16 \text{ 或 } 20(\text{即 } L/2 \text{ 是偶数}); \\ -\dfrac{L}{2}, & L = 10 \text{ 或 } 18; \\ -\dfrac{L}{2}+2, & L = 14. \end{cases} \tag{112e}$$

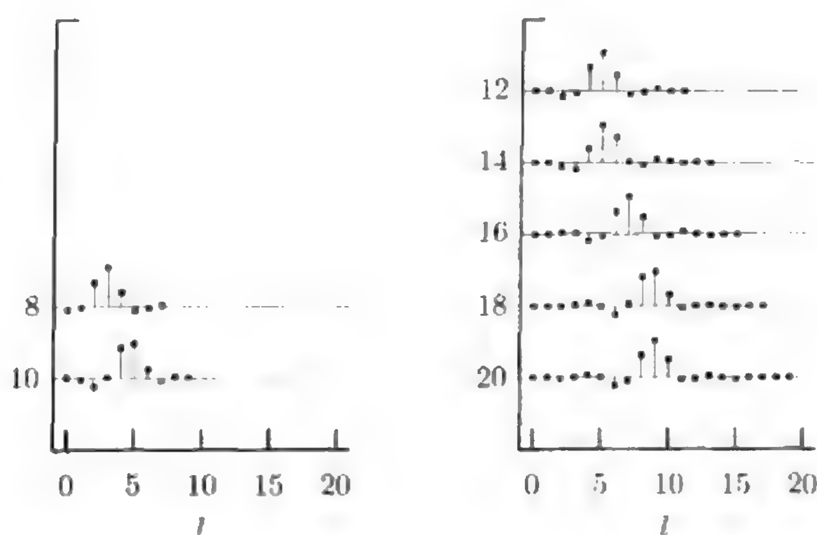


图 113a Daubechies 最接近对称的尺度滤波器 $\{g_l^{(la)}\}$，$L=8$，10，…，20(基于 Daubechies，1992，p. 198，表 6.3，有些变动，参见本节评论与扩展[1])

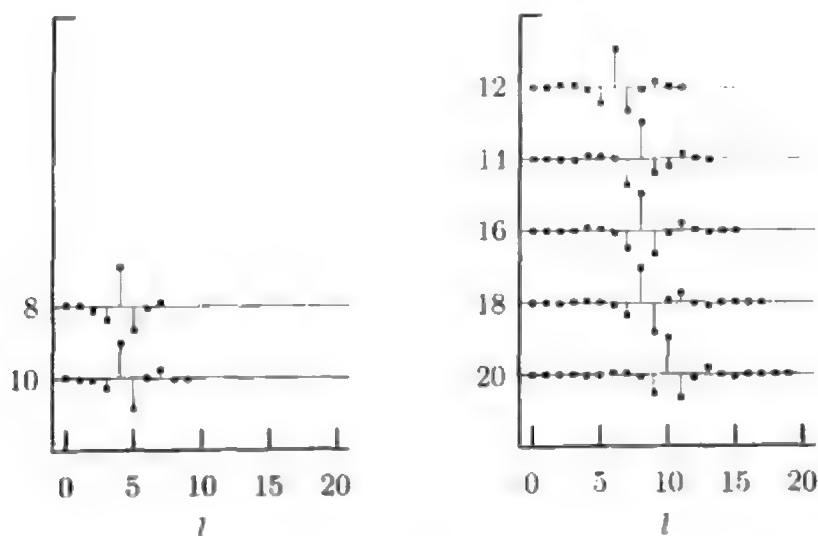


图 113b 相应于图 113a 给出的最接近对称的 Daubechies 尺度滤波器的小波滤波器

注意 ν 总是为负，因此我们可以通过前移 $|\nu|$ 个单位获得近似零相位的尺度滤波器. 类似地，因为 $-(L-1+\nu)$ 总是为负，我们可以通过前移 $|L-1+\nu|$ 个单位得到零相位的小波滤波器.

现在考虑在更高尺度 $\lambda_j=2^j$ 和 $\tau_j=2^{j-1}$ $(j=2, 3, \cdots)$ 的尺度滤波器 $\{g_{j,l}\}$ 与小波滤波器 $\{h_{j,l}\}$ 的相位函数. 在 4.6 节我们已经指出，这些滤波器的传递函数 $G_j(\cdot)$ 和 $H_j(\cdot)$ 满足

$$G_j(f) = \prod_{l=0}^{j-1} G(2^l f), \quad H_j(f) = H(2^{j-1} f) G_{j-1}(f).$$

练习[114]　令 $\theta_j^{(G)}(\cdot)$ 和 $\theta_j^{(H)}(\cdot)$ 是相应于 $G_j(\cdot)$ 和 $H_j(\cdot)$ 的相位函数. 利用上述关系及式(112d)，证明

$$\theta_j^{(G)}(f) \approx 2\pi f \nu_j^{(G)}, \quad 其中\ \nu_j^{(G)} \equiv (2^j - 1)\nu \tag{114a}$$

和

$$\theta_j^{(H)}(f) \approx 2\pi f \nu_j^{(H)}, \quad 其中\ \nu_j^{(H)} \equiv -[2^{j-1}(L-1)+\nu]. \tag{114b}$$

◁

由于滤波器 $\{g_{j,l}\}$ 和 $\{h_{j,l}\}$ 的长 L_j 可以表示成 $L_j=(2^j-1)(L-1)+1$(参见练习[96])，我们可以将上式重写为

$$\nu_j^{(G)} = \frac{L_j-1}{L-1}\nu = \begin{cases} -\dfrac{(L_j-1)(L-2)}{2(L-1)}, & L/2\ 是偶数; \\ -\dfrac{(L_j-1)L}{2(L-1)}, & L=10\ 或\ 18; \\ -\dfrac{(L_j-1)(L-4)}{2(L-1)}, & L=14 \end{cases}$$

和

$$\nu_j^{(H)} = -\left(\frac{L_j}{2}+\frac{L}{2}+\nu-1\right) = \begin{cases} -\dfrac{L_j}{2}, & L/2\ 是偶数; \\ -\dfrac{L_j}{2}+1, & L=10\ 或\ 18; \\ -\dfrac{L_j}{2}-1, & L=14. \end{cases} \tag{114c}$$

对于小波滤波器 $\{h_{j,l}\}$，$\nu_j^{(H)}$ 总是为负，因此我们通过将 $\{h_{j,l}\}$ 前移 $|\nu_j^{(H)}|$ 个单位得到了最接近零相位的滤波器. 这些滤波器的相位函数在图115中给出了，频率范围(也是其标称通带范围)为 $[1/2^{j+1}, 1/2^j]$，$j=1, 2, 3$ 和 4. 很明显，当 $L/2$ 为偶数时这些滤波器更好地逼近了零相位滤波器(注意 D(4)滤波器的相位特性相当于 $L/2$ 的最接近对称滤波器).

在实际应用中我们用位移 $\nu_j^{(H)}$ 和 $\nu_{J_0}^{(G)}$——为了绘图——使成为基于 Daubechies 最接近对称小波滤波器之一的 J_0 层部分离散小波变换小波的系数和尺度系数 $\mathbf{W}_j$ 和 $\mathbf{V}_{J_0}$. 注意

$$W_{j,t} = 2^{j/2}\widetilde{W}_{j,2^j(t+1)-1}, \quad t=0,\cdots,N_j-1 \equiv \frac{N}{2^j}-1,$$

其中，对于 $t=0, \cdots, N-1$，

$$2^{j/2}\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} h_{j,l} X_{t-l \bmod N}.$$

现在需要将第 j 层小波滤波器 $\{h_{j,l}\}$ 前移 $|\nu_j^{(H)}|$ 个单位以达到近似的零相位. 或者等价地，我们将滤波器输出 $\{\widetilde{W}_{j,t}\}$ 前移 $|\nu_j^{(H)}|$ 个单位. 如果令 $\widetilde{\mathbf{W}}_j$ 是 N 维向量，其元素为 $\widetilde{W}_{j,t}$，注意 $\nu_j^{(H)}<0$，因此我们用 $\nu_j^{(H)}=-|\nu_j^{(H)}|$ 以保持方向，这个前移的输出可以用向量形式表示 $\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j$，其中第 t 个元素和 X_t 相关联；即我们将输

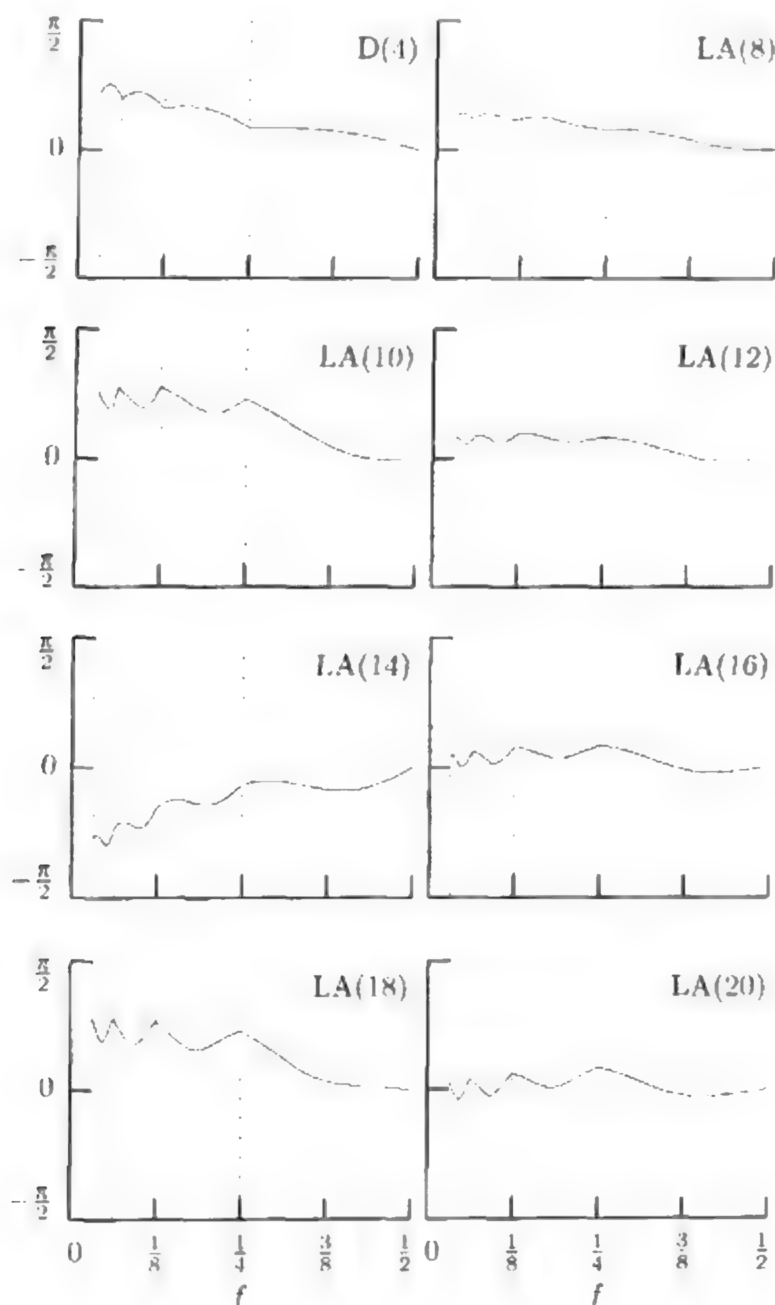


图 115　平移后的最接近对称小波滤波器$\{h_{j,l}\}$($j=1, 2, 3, 4$)的相位函数（同时也给出了D(4)对应的相位函数，其中平移因子为$\nu=-1$）．相位函数$\{h_{j,l}\}$只绘出了在区间$[1/2^{j+1}, 1/2^j]$上的部分——这些区间由垂直的点划线标出

出$\widetilde{W}_{j,(t+|\nu_j^{(H)}|) \bmod N}$和输入$X_t$相关联，或者等价地，将$\widetilde{W}_{j,t}$和$X_{(t-|\nu_j^{(H)}|) \bmod N}$相关联．最后，由于$W_{j,t}=2^{j/2}\widetilde{W}_{j,2^j(t+1)-1}$，我们看到$W_{j,t}$和$X_{(2^j(t+1)-1-|\nu_j^{(H)}|) \bmod N}$相关联．类似的论证表明，尺度系数$V_{J_0,t}$与$X_{(2^{J_0}(t+1)-1-|\nu_{J_0}^{(G)}|) \bmod N}$相关联.

现在假设时间序列的第t个值X_t实际上和时间$t_0+t\Delta t$相关联，其中t_0(单位可以是秒)是采样起始时间，Δt是两个采样之间的时间间隔(和t_0的单位相同)．因此，时间序列的实际图形是纵坐标为X_t，横坐标为$t_0+t\Delta t$，

$t=0,\cdots,N-1$. 上面的论证说明，为了绘出时间序列的真实图像，我们需要绘出横坐标为 $t_0+(2^j(t+1)-1-|\nu_j^{(H)}| \bmod N)\Delta t(t=0,\cdots,N_j-1)$的序列 $W_{j,t}$ 以及横坐标为 $t_0+(2^{J_0}(t+1)-1-|\nu_{J_0}^{(G)}| \bmod N)\Delta t(t=0,\cdots,N_{J_0}-1)$的序列 $V_{J_0,t}$.

练习[116] 作为这种调整过程的具体说明，考虑一个时间序列$\{X_t: t=0,\cdots,127\}$，长度 $N=128$，使 $t=63$ 时，$X_{63}=1$，t 为其他值时，$X_t=0$. 假定 $t_0=17$ 且 $\Delta t=1$，于是 X_t 和实际时间 $t+17$ 相关联. 计算关于这个序列$\{X_t\}$的层数 $J_0=3$ 的 LA(8)的部分离散小波变换，并且计算平移 $\nu_j^{(H)}(j=1,2,3)$和 $\nu_3^{(G)}$. 对于 $j=1,2,3$ 绘出以 $t_0+(2^j(t+1)-1-|\nu_j^{(H)}| \bmod N)$为横坐标$(t=0,\cdots N_j-1\equiv N/2^j-1)$的序列 $\mathbf{W}_j$ 的元素 $W_{j,t}$的图形；对 $\mathbf{V}_3$ 的元素 $V_{3,t}$同样绘出图形. 说说看绘出来的$\{X_t\}$的小波系数和尺度系数通过直线连接起来后看起来光滑程度怎么样. ◁

正如上述练习所表明的，$\mathbf{W}_j$ 的元素顺序和实际画出来的顺序是不同的. 然而通过循环平移 $\mathbf{W}_j$ 一定的位数，即$\mathcal{T}^{-\gamma_j^{(H)}}\mathbf{W}_j$，其元素顺序反映了绘出来的 $\mathbf{W}_j$ 的元素顺序(如何确定这个整数 $\gamma_j^{(H)}$ 将在 4.11 讨论). 在后面的图 127 和图 138 中，我们给出了以实际时间为横坐标的 $\mathbf{W}_j$ 的图形，我们将这些图形认为是 $\mathcal{T}^{-\gamma_j^{(H)}}\mathbf{W}_j$ 而不是 $\mathbf{W}_j$，主要是为了强调我们不是简单地绘以$t=0,\cdots,N_j-1$ 为横坐标的 $W_{j,t}$(对 $\mathbf{V}_{J_0}$ 也是一样).

4.8 节的关键结论与定义

Daubechies 尺度滤波器$\{g_l: l=0,\cdots,L-1\}$是具有平方增益函数为

$$\mathcal{G}^{(D)}(f)\equiv 2\cos^L(\pi f)\sum_{l=0}^{\frac{L}{2}-1}\binom{\frac{L}{2}-1+l}{l}\sin^{2l}(\pi f)$$

的任意滤波器，其中 L 是一个正偶数. 相应的 Daubechies 小波滤波器$\{h_l: l=0,\cdots,L-1\}$可以用正交镜像滤波器关系从式(75b)得到，并且其平方增益函数为

$$\mathcal{H}^{(D)}(f)\equiv 2\sin^L(\pi f)\sum_{l=0}^{\frac{L}{2}-1}\binom{\frac{L}{2}-1+l}{l}\cos^{2l}(\pi f).$$

上式可以重新表示为$\mathcal{H}^{(D)}(f)=\mathcal{D}^{\frac{L}{2}}(f)\mathcal{A}_L(f)$，其中$\mathcal{D}(\cdot)$是差分滤波器$\{1,-1\}$的平方增益函数，而$\mathcal{A}_L(\cdot)$是某个低通滤波器的平方增益函数(见图 107).

对于固定的 L，平方增益函数$\mathcal{G}^{(D)}(\cdot)$不能唯一决定尺度滤波器. 所有以$\mathcal{G}^{(D)}(\cdot)$为平方增益函数的滤波器$\{g_l\}$可以通过"谱分解"得到. 我们的目的是找到传递函数 $G(\cdot)$满足 $|G(f)|^2=\mathcal{G}^{(D)}(\cdot)$. 从某种意义上说，就是找出$\mathcal{G}^{(D)}(\cdot)$的"平方根". 举个例子，当 $L=2$ 时，存在两种可能的分解，一个是哈尔尺度滤波

器$\left\{\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\}$，另一个是$\left\{-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right\}$. 当 $L=4$ 时，我们得到式(75d)中的D(4)尺度滤波器，另外三个滤波器只在方向和符号上有所不同. 对于$L\geqslant 6$，就有很多不同的滤波器了，相互之间的差异也没有上面两种情况那么简单，因此我们需要另外的标准来从多种分解方式中选择其中一种. 其中一种标准是极小相位，这种标准导致了最小延迟滤波器，这个标准选择尺度滤波器$\{g_l^{(\mathrm{ep})}\}$满足平方增益函数为$\mathcal{G}^{(\mathrm{D})}(\cdot)$，使得

$$\sum_{l=0}^{m} g_l^2 \leqslant \sum_{l=0}^{m}[g_l^{(\mathrm{ep})}]^2, \quad m=0,\cdots,L-1,$$

其中$\{g_l\}$是任意另外满足平方增益函数为$\mathcal{G}^{(\mathrm{D})}(\cdot)$的滤波器. 我们记满足这个标准的滤波器为 D($L$)尺度滤波器，$L=2$，4，…. 这些滤波器在图 108a 中给出了(相应的小波滤波器在图 108b 中给出).

另一个标准是选择尺度滤波器$\{g_l^{(\mathrm{la})}\}$，其传递函数为$G(f)=[\mathcal{G}^{(\mathrm{D})}(f)]^{1/2}e^{i\theta^{(G)}(f)}$是使相位函数$\theta^{(G)}(\cdot)$最接近线性相位(在式(112a)的意义上)的滤波器. 我们记这个标准下选择的滤波器为 LA(L)尺度滤波器，$L=8$，10，…，其中“LA”是 least asymmetric(最接近对称)的缩写(这个名称是合适的，因为真正的线性相位都是关于它的中点对称的). 这些滤波器在图 113a 中给出了(最接近对称小波滤波器在图 113b 中给出). 最接近对称尺度滤波器的相位函数是$\theta^{(G)}(f)\approx 2\pi f\nu$，其中$\nu$依赖于式(112e)中的 L.

4.8 节的评论与扩展

[1]Daubechies(1992) 的表 6.3 给出了宽 L 是从 8 到 20 的偶数的最接近对称尺度滤波器系数，且系数都规范化为$\sum_l g_l=2$，而不是我们这里用的$\sum_l g_l=\sqrt{2}$. 计算表明式(112c)中的ν对$L=8$，12，18 和 20 为奇数，对$L=10$，14，16 则为偶数. 因此，LA(8)、LA(12)、LA(18)和 LA(20)四个尺度滤波器系数具有近似线性相位，而另外三个滤波器则称为是具有近似广义线性相位，广义线性相位定义为：线性相位加上一个非零的中间项(Oppenheim and Schafer, 1989)；这里非零中间项可以取成$\theta^{(H)}(0)=\pi$. 下面这个练习的结果可以用来消除这个非零中间项.

练习[117] 证明如果滤波器$\{g_0, g_1, \cdots, g_{L-2}, g_{L-1}\}$具有传递函数$G(\cdot)$，则反序的滤波器$\{g_{L-1}, g_{L-2}, \cdots, g_1, g_0\}$具有传递函数$\exp[-i2\pi f(L-1)]G^*(f)$. ◁

于是，如果 Daubechies 尺度滤波器的系数反序之后，相应的反序的尺度滤波器的传递函数是$\exp[-i2\pi f(L-1)]G^*(f)$，因此得到的相位函数为

$$-\theta^{(G)}(f)-2\pi f(L-1)\approx 2\pi f(-\nu_b-[L-1])\equiv 2\pi f\nu,$$

其中ν_b在反序之前是偶数，反序之后的ν就是奇数了. 我们在图 113a 对于$L=10$，14，16 给出了图示(对于$L=16$其值在表 109 中给出了). 因此$\{g_l\}$是从

表 118

L	$\lvert\nu\rvert-e\{g_l\}$	L	$\lvert\nu\rvert-e\{g_l\}$
8	0.153 6	14	−0.161 5
10	0.450 0	16	0.154 6
12	0.154 3	18	0.447 1
		20	0.154 7

注：Daubechies 最接近对称的小波滤波器的前移因子 $\lvert p_j^{(H)}\rvert$ 和 $\lvert\nu_j^{(H)}\rvert$ 的比较．由于 $\lvert p_j^{(H)}\rvert-\lvert\nu_j^{(H)}\rvert=\lvert\nu\rvert-e\{g_l\}$ 对所有的指标 j 成立，上面的表说明当四舍五入到一个整数后这两个前移是相同的．为了比较相应的尺度滤波器的前移，上面的值必须乘以 2^j-1，因此当 j 增大时这个前移是发散的．

Daubechies(1992)给出的尺度滤波器经过一个反序得到的．

[2]Wickerhauser(1994，pp. 171 和 341)和 Hess-Nielsen and Wickerhauser(1996)给出了前移量 $\lvert p_j^{(G)}\rvert$ 与 $\lvert p_j^{(H)}\rvert$，它们可以同时对极小相位和最接近对称滤波器同时适用(也包括其他滤波器)．这些前移量为

$$\lvert p_j^{(G)}\rvert=(2^j-1)e\{g_l\},\quad \lvert p_j^{(H)}\rvert=2^{j-1}(e\{g_l\}+e\{h_l\})-e\{g_l\},$$

其中

$$e\{a_l\}\equiv\frac{\sum_{l=1}^{L-1}la_l^2}{\sum_{l=0}^{L-1}a_l^2}$$

表示滤波器 $\{a_l: l=0,\cdots,L-1\}$ 的“能量中心”．

练习[118] 因为知道尺度滤波器 $\{g_l\}$ 和小波滤波器 $\{h_l\}$ 具有单位能量且满足式(75a)的正交镜像滤波器关系，证明必有 $e\{g_l\}+e\{h_l\}=L-1$. ◁

对于所有小波滤波器和尺度滤波器，我们有

$$\lvert p_j^{(G)}\rvert=(2^j-1)e\{g_l\},\quad \lvert p_j^{(H)}\rvert=2^{j-1}(L-1)-e\{g_l\}.$$

我们能够把上述滤波器与通过使用式(114a)与(114b)的右部分得出的最小对称滤波器进行对比：

$$\lvert\nu_j^{(G)}\rvert=(2^j-1)\lvert\nu\rvert,\quad \lvert\nu_j^{(H)}\rvert=2^{j-1}(L-1)-\lvert\nu\rvert,$$

其中，对于不同的滤波器长度 L，ν 由式(112e)给出．于是我们看到

$$\lvert\nu_j^{(G)}\rvert-\lvert p_j^{(G)}\rvert=(2^j-1)(\lvert\nu\rvert-e\{g_l\}),\quad \lvert p_j^{(H)}\rvert-\lvert\nu_j^{(H)}\rvert=\lvert\nu\rvert-e\{g_l\}.$$

对于 $L=8$ 到 20，差 $\lvert\nu\rvert-e\{g_l\}$ 的值在表 118 中给出了，从表中我们可以看出其绝对值总是小于 0.5 的．这说明，当 $\lvert p_j^{(H)}\rvert$ 四舍五入成一个整数时，小波系数的前移同样由 $\lvert\nu_j^{(H)}\rvert$ 或 $\lvert p_j^{(H)}\rvert$ 决定．另一方面，当 j 增大时尺度系数的前移变得分散(我们将在第 6 章考虑离散小波包变换时继续讨论滤波器前移)．

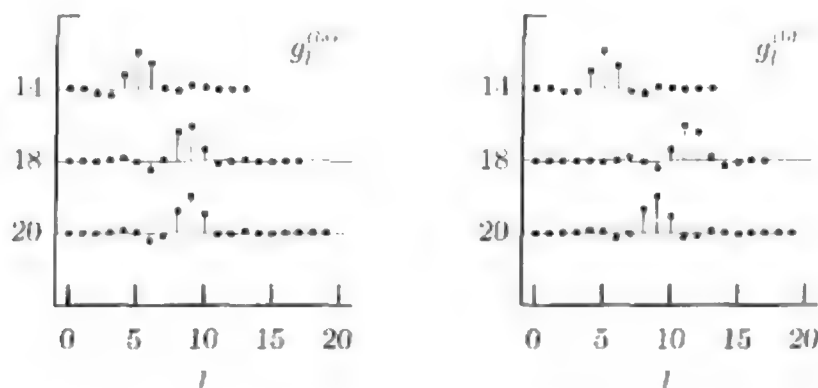


图 119　最接近对称的尺度滤波器$\{g_l^{(la)}\}$(左图)和最优局部的尺度滤波器$\{g_l^{(bl)}\}$(右图)，$L=14$，18，20

[3]Doroslovacki(1998)对 Daubechies 尺度滤波器的平方增益函数给出了"最优局部"(BL)的分解——这个方案试图进一步改善最接近对称的滤波器. 最接近对称分解是对线性相位的相位函数在频域上最大偏移量进行极小化得到的；然而，由于尺度滤波器是低通的，高频部分的相位其实不是很重要，更重要的是低频部分的相位. 在 Doroslovacki 的工作中他用了一种新的度量来衡量滤波器和线性相位的偏离程度，这个度量对低频部分乘一个较大的权值，对高频部分乘较小的权值. 对于 $L=8$，10，12 和 16，这种新的度量得到和最接近对称一样的分解. 对于 $L=14$，18 和 20，这种最优局部分解和最接近对称分解得到的结果就不同了. 这些新的尺度滤波器在图 119 的右图中给出，而对应的最接近对称尺度滤波器则在左图中给出(注意到我们这里给出的 $L=14$ 和 18 的最优局部和 Doroslovacki 在 1998 年给出的系数差了一个反序，同样，我们的 LA(10)、LA(14)和 LA(16)也和 Daubechies 给出的差了一个反序).

如果对 $L=14$ 取 $\nu=-5$；对 $L=18$ 取 $\nu=-11$；对 $L=20$ 取 $\nu=-9$，通过计算表明这时这些滤波器的相位函数 $\theta^{(G)}(\cdot)$ 近似有 $\theta^{(G)}(f)\approx 2\pi f\nu$. 图 120 右面一列图给出了这些最优局部小波滤波器$\{h_{j,l}\}$，$j=1,2,3,4$ 在标称通带中的精确的相位函数图像. 而左面一列图中给出了相应的最接近对称小波滤波器的相位函数图像(这些是从图 115 提取的). 因为理想情况下由于这些曲线应该非常接近于零，我们看到 BL(14)很明显比 LA(14)更接近于线性相位；然而，对于 BL(18)和 BL(20)滤波器情况就不是这样了，这两个滤波器看起来都要比相应的 LA(18)和 LA(20)差一点(由于 BL(18)和 LA(18)的相位函数符号相反，因此我们同时绘出了负的 LA(18)的相位函数——即那条较粗的线，这样更容易比较).

Taswell(2000)讨论了另一些分解的标准和系统的求解相应滤波器的方法.

[4]让我们更仔细地讨论一下最接近对称小波滤波器的相位性质. Koopmans (1974，pp. 95～96)定义了一个时间平移函数，可以将相位函数 $\theta(\cdot)$ 给出的角偏移根据指标 t 进行平移：

$$\tau(f)\equiv\frac{\theta(f)}{2\pi f},\quad f\neq 0\,.$$

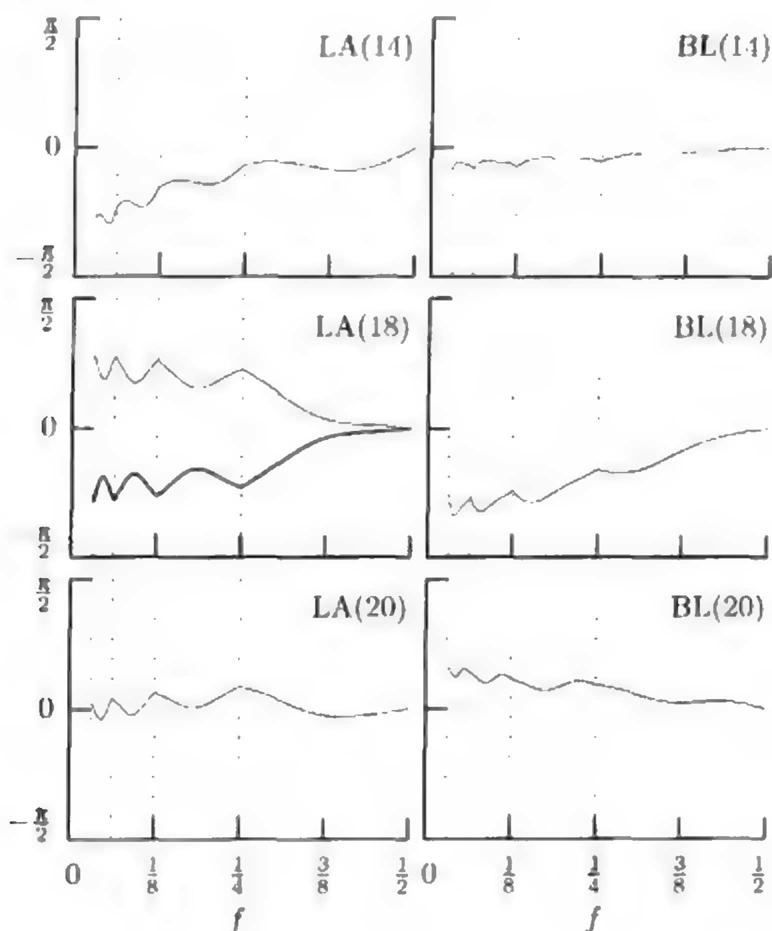


图 120 平移后的最接近对称的小波滤波器和最优局部的小波滤波器的精确的相位函数(参见图 115 及其说明)

例如，设 $\theta(f)=2\pi f\nu$，ν 为某个整数；即 $\theta(\cdot)$是线性相位滤波器的相位函数. 于是时间平移函数为 $\tau(f)=\nu$，即我们为了得到零相位滤波器而需要将输出前移的量.

在图 115 中我们看到最接近对称滤波器被前移之后比较接近于零相位. 如果前移后的滤波器是真正的零相位滤波器，图上所有的曲线都成了水平线，且纵坐标恒为零. 由于只是近似逼近于水平线，因此就可以将角偏移用指标 t 来表示. 图 121 给出了四舍五入后的时间平移函数. 注意对于单位尺度$\left(\text{即频带为}\left[\frac{1}{4},\ \frac{1}{2}\right]\right)$，四舍五入后的平移函数总为零——因为前移量总是使得滤波器在单位尺度上更接近于零相位. 随着尺度的增加，四舍五入之后的时间平移函数的非零开始出现，特别是图 121 左图中的值. 这些结果意味着我们实际上可以通过额外的由通带内时间平移函数的平均四舍五入后得到的前移获得在$\left[\frac{1}{2^{j+1}},\ \frac{1}{2^j}\right]$上一个更好的逼近于线性相位的滤波器.

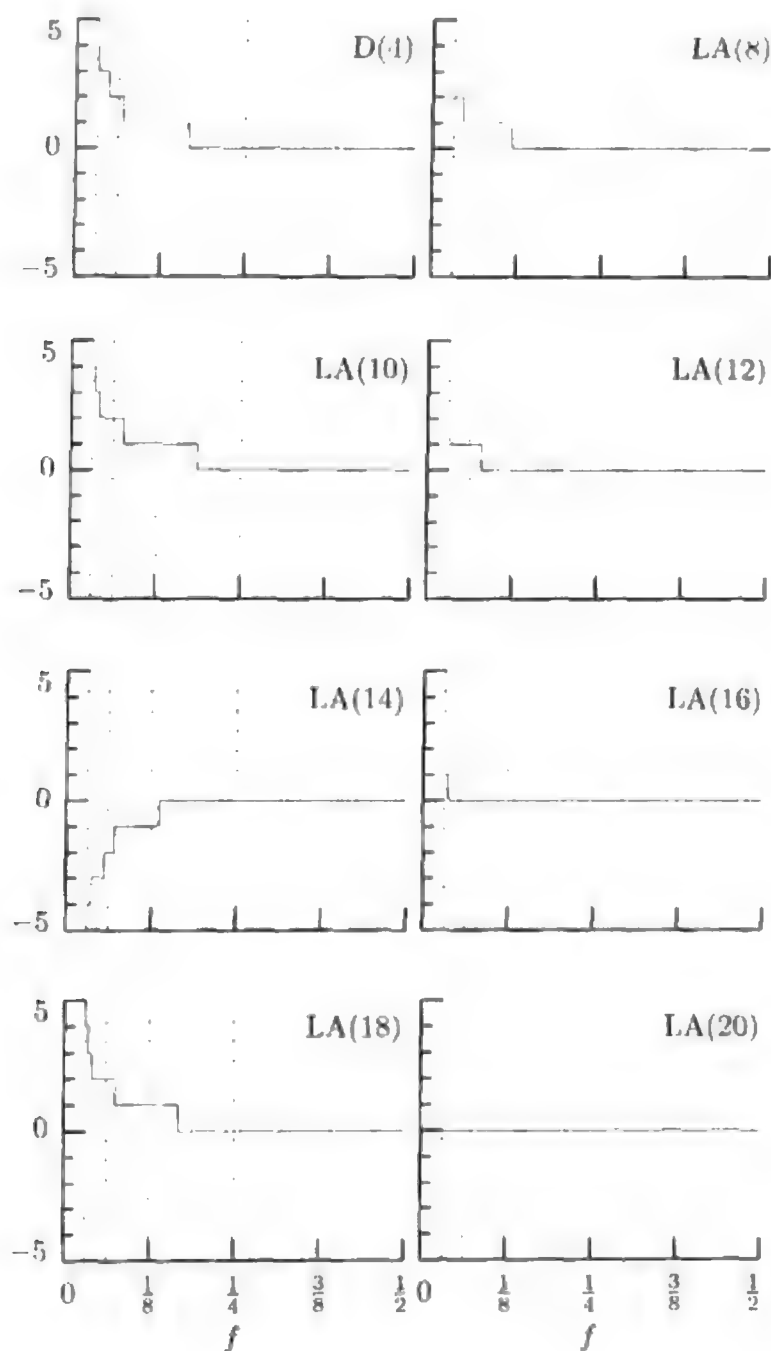


图 121 相应于图 115 给出的相位函数的时间平移函数(四舍五入为整数)

例如，这时我们需要将 $j=3$ 层的 LA(8)滤波器的输出前移一个单位.

[5]让我们仔细考虑 Daubechies 小波滤波器的平方增益函数 $\mathcal{H}_j(\cdot)$. 图 122 对 $L=2, 4, \cdots, 14$ 和尺度 $j=1, 2$ 和 3 给出了这些函数的图像. 垂直的细线表示通带$(1/2^{j+1}, 1/2^j]$，函数是在分贝尺度(dB)上绘的；即我们绘的是以 f 为横坐标，$10 \cdot \log_{10}(\mathcal{H}_j(f))$为纵坐标的图像. 坐标刻度间隔为 10dB(即单位长度的大小). 注意随着 L 的增加，通带逼近于$(1/2^{j+1}, 1/2^j]$，注意这些滤波器随着 L 增加，其通带也逼近于理想滤波器，实际上这种逼近不是偶然的，但是一旦我们给出哈尔滤波器，在由 L 到 $L+2$ 的改进中不是引人注目的. 因为

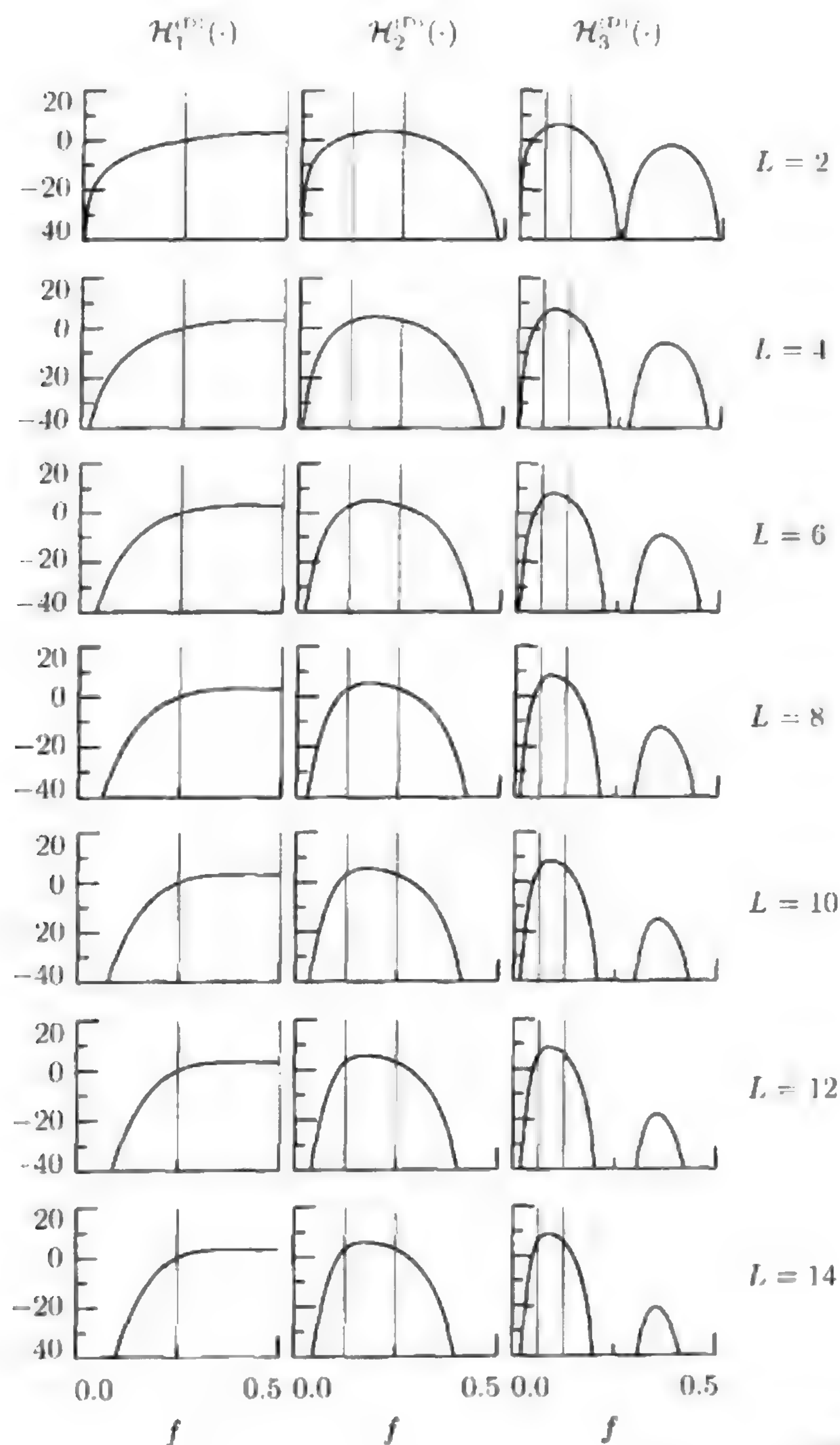


图 122 宽为 $L=2, 4, \cdots, 14$(从上到下)的 Daubechies 小波滤波器的平方增益函数 $\mathcal{H}_j^{(D)}(\cdot)$，$j=1, 2, 3$(分别是左列图、中列图和右列图)．每幅图中的垂直细线标记了滤波器的通带范围．纵轴坐标刻度为分贝(即我们绘的是纵轴为 $10\cdot\log_{10}(\mathcal{H}_j^{(D)}(f))$ 横轴为 f 的图)

$\mathcal{H}_j(f)=\mathcal{H}(2^{j-1}f)\mathcal{G}_{j-1}(f)$ 且对于小的 f 来说有 $\mathcal{G}_{j-1}(f)\approx 2^{j-1}$，$\mathcal{H}_j(\cdot)$ 在 $f=1/2^{j+1}$ 处衰减到零的速度主要由 $\mathcal{H}(\cdot)$ 的衰减速度控制．由式(105b)，这个衰减速度又是由其中 $L/2$ 阶差分的阶数控制，因此对于小的 f，我们有 $\mathcal{H}_j(f)\propto|f^L|$．注意 $\mathcal{H}_3(\cdot)$ 的主通带右面的旁瓣也很明显．当 j 增大时，这些右面的

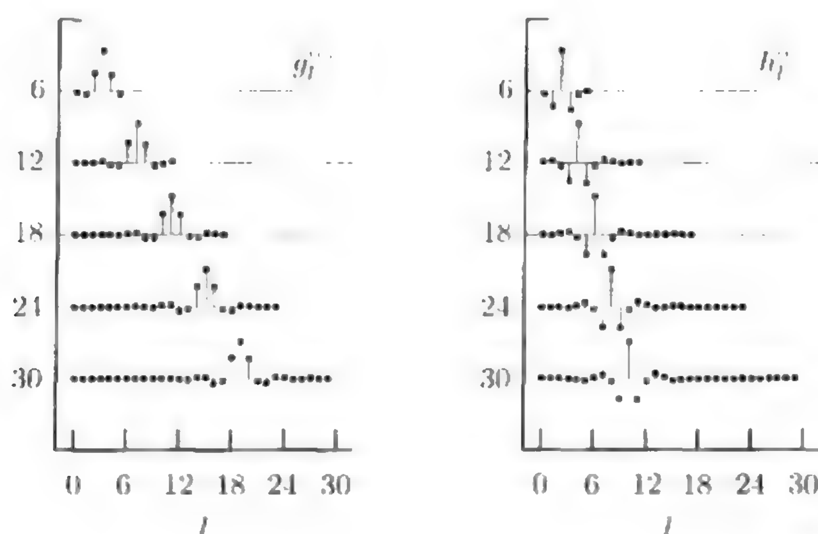


图 123　coiflet 尺度滤波器$\{g_l^{(c)}\}$(左图)和相应的小波滤波器$\{h_l^{(c)}\}$(右图)，$L=6$，12，18，24，30(Daubechies，1992，p. 261，表 8. 1)

旁瓣越来越多：对于 $j\geq3$，$\mathcal{H}_j(\cdot)$有 $2^{j-2}-1$ 个这样的旁瓣(这主要时由于上式中的$\mathcal{G}(2^l f)$是一个周期为 $1/2^l$ 的函数). 就像我们将要在 8. 9 节看到的那样，$\mathcal{H}_j(\cdot)$对理想带通滤波器的逼近对我们解释小波方差有重要影响(参见第 8 章).

4. 9　Coiflet 小波滤波器和尺度滤波器：形式和相位

重要的是认识到，Daubechies 小波滤波器不是可以导出离散小波变换且可以解释为对序列加权平均进行广义差分的唯一的小波滤波器. 这样的小波滤波器的第二个例子是"coiflets"，这个词是 Daubechies(1992，8. 2 节) 首先提出来的，因为是 Coifman 先给出这种构造的想法. 虽然我们是通过平方增益函数来定义小波滤波器的(式(105a)和(105b))，但实际上我们是通过加上完全由尺度滤波器决定的小波函数的消失矩条件来得到这些条件的(参见 11. 9 节). 构造 coiflets 的想法就是对尺度函数也加上消失矩条件. 这种构造最终得到了一组有着很好的线性相位性质的小波滤波器.

图 123 对 $L=6$，12，18，24 和 30 给出了 coiflet 尺度滤波器$\{g_l^{(c)}\}$和相应的小波滤波器$\{h_l^{(c)}\}$，它们最早是由 Daubechies(1992，p. 261，表 8. 1)给出——我们以后用 C(L)表示宽为 L 的 coiflet 滤波器. 表 109 给出了尺度滤波器 C(6)的系数. 注意我们这里给出的系数和 Daubechies 的系数是差一个反序的，原因和前面给出的 LA(10)、LA(14)和 LA(16)的尺度系数是反序的原因是一样的(参见 4. 8 节的评论与扩展第 1 条).

$\{h_l^{(c)}\}$的平方增益函数可以写为

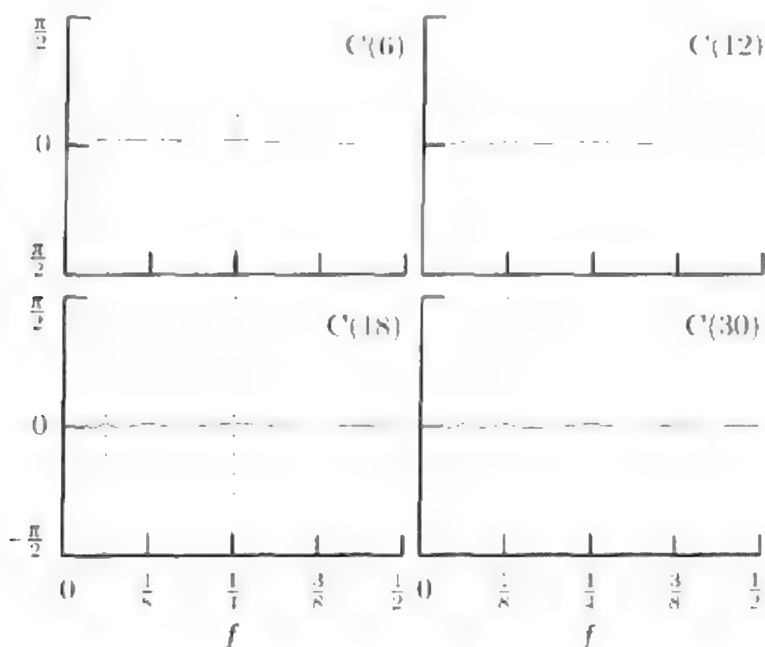


图 124 平移 cioflet 小波滤波器后的精确相位函数(参见图 115 及其说明，详情参见正文)

$$\mathcal{H}^{(c)}(f) = \mathcal{D}^{\frac{L}{3}}(f)\left[\sum_{l=0}^{\frac{L}{6}-1}\binom{\frac{L}{6}-1+l}{l}\cos^{2l}(\pi f)+\cos^{\frac{L}{3}}(\pi f)F(f)\right]^2,$$

其中 $F(\cdot)$是三角多项式，满足式(69d)，即 $\mathcal{H}^{(c)}(f)+\mathcal{H}^{(c)}\left(f+\frac{1}{2}\right)=2$ 对所有的 f 成立. 注意 Daubechies 小波滤波器的宽度为 L，且包含了 $L/2$ 个差分运算，而 coiflet 滤波器包含了 $L/3$ 个这样的差分，因此包含加权平均的项将更多. 计算表明$\{g_l^{(c)}\}$的相位函数近似为 $2\pi f\nu$，而 $\nu=-\frac{2L}{3}+1$(注意当 $L=6$，12，18，24，30 时，这个数总是奇数). 从式(114a)和(114b)可以得到平移因子

$$\nu_j^{(G)}=-\frac{(L_j-1)(2L-3)}{3(L-1)},\quad \nu_j^{(H)}=-\frac{L_j}{2}+\frac{L}{6}. \tag{124}$$

用图 115 中对最接近对称小波滤波器同样的方法，图 124 给出了对应于 $L=6$，12，18 和 30 的平移后的 coiflet 小波滤波器精确的相位函数($L=24$ 没有给出，因为它与 $L=18$ 和 $L=30$ 很类似). 对比图 115 和图 124 可以发现 coiflets 是滤波器比最接近对称滤波器更好的逼近零相位滤波器.

图 125 给出了基于 C(6)的尺度滤波器$\{g_{j,l}\}$和小波滤波器$\{h_{j,l}\}$的图形，$j=1,\cdots,7$. (图 98a 和 98b 给出了 D(4)和 LA(8)的类似的图形.)注意随着 j 的增加，$\{g_{j,l}\}$收敛于近似的加权方案三角形，只是在顶上有一个尖端. 类似于鲨鱼鳍的D(4)滤波器能引入多分辨分析的人工痕迹，因此 C(6)也有可能会这样. 这个滤波器具有比 LA(8)更好的相位性质，其三角形外形却不能像 LA(8)那样

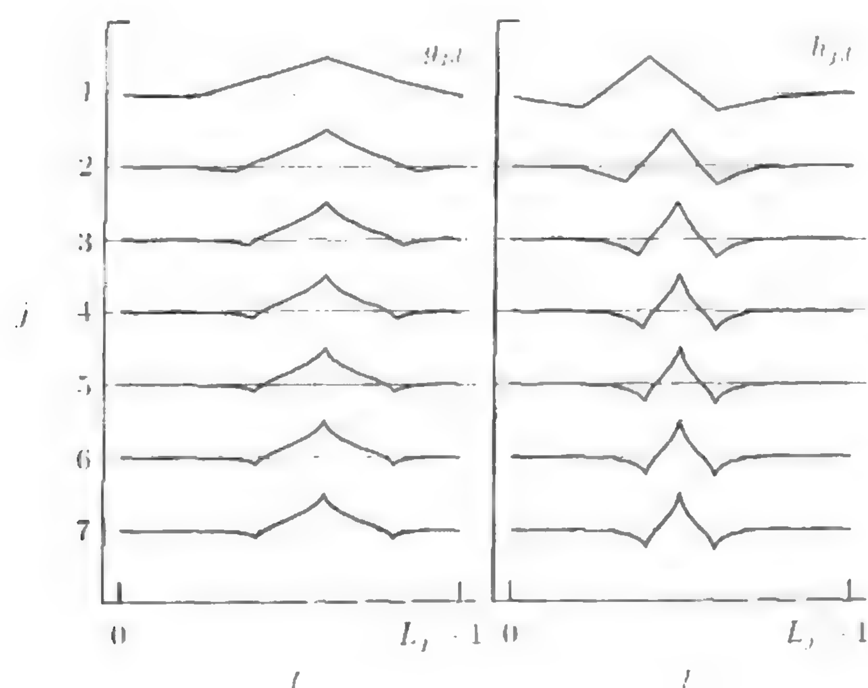


图 125　在尺度指标 $j=1, 2, \cdots, 7$ 时的 C(6) 尺度滤波器 $\{g_{j,l}\}$ 和小波滤波器 $\{h_{j,l}\}$. (对于 D(4) 和 LA(8) 滤波器的类似的图见图 98a 和 98b)

可以很好地刻画时间序列的特征. 因此 C(6)和 LA(8)滤波器, 一个具有好的相位性质, 另一个不会引入人工痕迹, 我们需要根据实际应用在这两者之间进行权衡. (另一个考虑是 C(6)滤波器用两个嵌入的差分运算而 LA(8)用四个嵌入的差分运算, 这意味着 LA(8)的平方增益函数在减小通带外的旁瓣的意义上可以更好地逼近于理想低通滤波器.)

4.10　例子: 心电图数据分析

作为使用离散小波变换分析实际数据的第一个例子, 我们现在考虑图 127 中的心电图(ECG)数据的时间序列. 这些数据是对心率不齐的病人在正常的时候测得的脉搏数据. 我们有 $N=2\,048$ 个观察数据, 其单位是毫伏, 并且这些数据是用 $\Delta t=1/180$s 的采样间隔得到的. 整个数据覆盖了 11.37 秒的时间长度(为了绘图的方便, 我们将 X_0 的数据的第一次采集时间定为 $t_0=0.31$s). 这个序列是进行多分辨分析的很好的候选者, 因为它的成分在不同尺度上都有分布. 例如, 对于大尺度(低频)的波动——即基线偏移——主要来自病人的呼吸, 而明显的小尺度(高频)的间歇(3～4 秒之间)性的波动显然是病人的其他动作. 这些波动都和心脏没关系, 但是心脏的跳动决定了这个序列的剩下的大部分特征. 在 0.7 秒时刻的长钉状波动是正常心跳的 R 波; 另一个在 R 波前面的小的尖峰被称为是 P 波, 在 P 波和 R 波后最宽的尖峰称为 T 波(这样的 PRT 复合体在图 127 右边底下有标记).

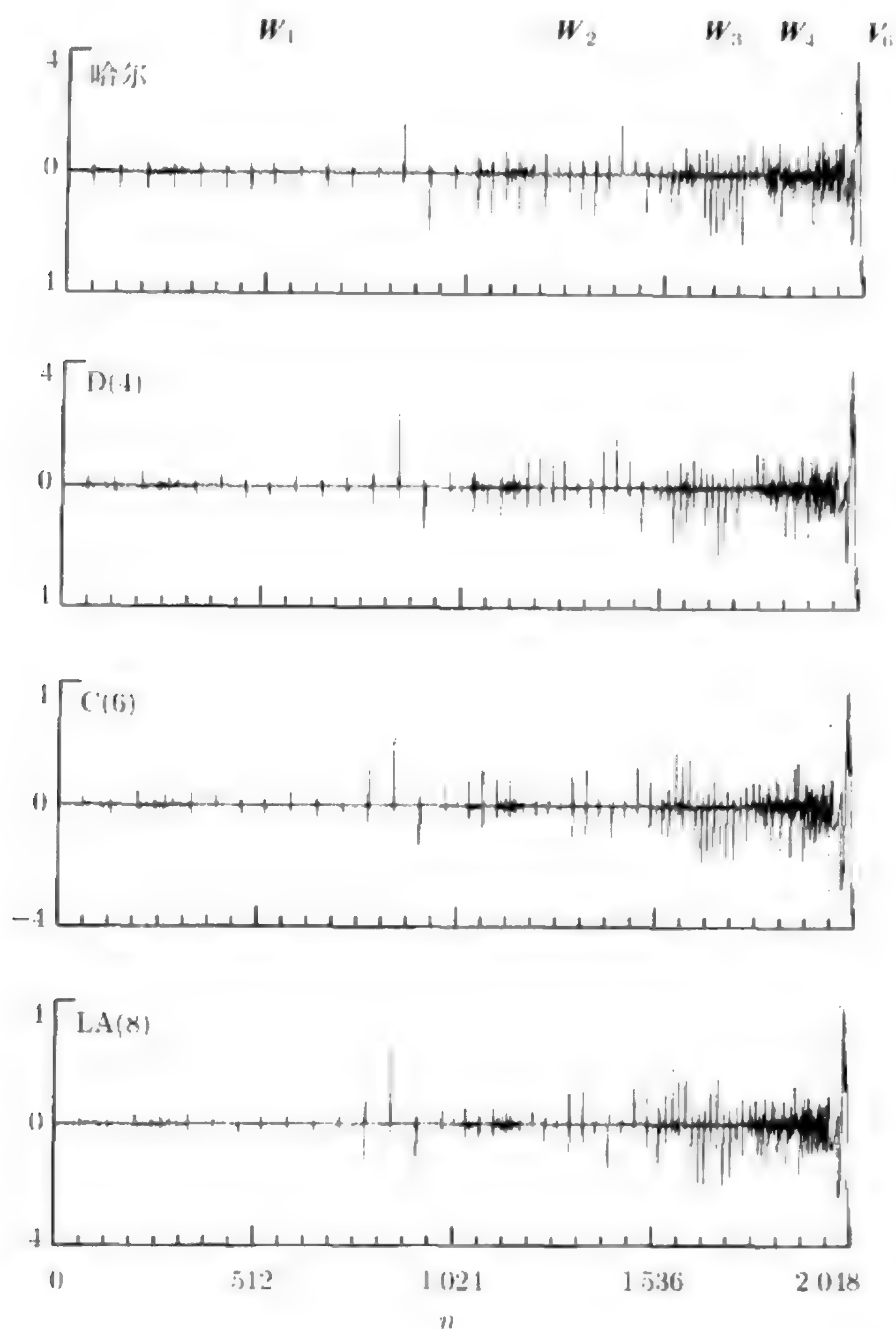


图 126 心电图时间序列的分别对应哈尔、D(4)、C(6) 和 LA(8) 小波的 $J_0=6$ 层部分离散小波变换系数 $\boldsymbol{W}$. 图上绘出了纵轴为 $\boldsymbol{W}$ 的元素 W_n，横轴为 $n=0,\cdots,N-1=2\,047$ 的图形. 6 条垂直点划线标记 $\boldsymbol{W}$ 的子向量，即 $\boldsymbol{W}_1,\cdots,\boldsymbol{W}_6$ 和 $\boldsymbol{V}_6$(由于空间不够，$\boldsymbol{W}_5$ 和 $\boldsymbol{W}_6$ 没有标记出来)

对 $J_0=6$，图 126 展示了基于哈尔、D(4)、C(6)和 LA(8)小波的部分离散小波变换系数向量 $\boldsymbol{W}$ 的元素. 垂直划线描绘了子向量 $\boldsymbol{W}_1,\cdots,\boldsymbol{W}_6$ 和 $\boldsymbol{V}_6$. 每个子向量 $\boldsymbol{W}_j$ 的小波系数的个数为 $2\,048/2^j$，其中 $\boldsymbol{V}_6$ 中有 $2\,048/2^6=32$ 个尺度系数，总共有 2 048 个系数. $\boldsymbol{W}$ 中的各元素的平方和与心电图时间序列的各元素的平方和相等(保持能量). 对于每个小波，最大的小波系数总在 $\boldsymbol{W}$ 的最后，也就是在

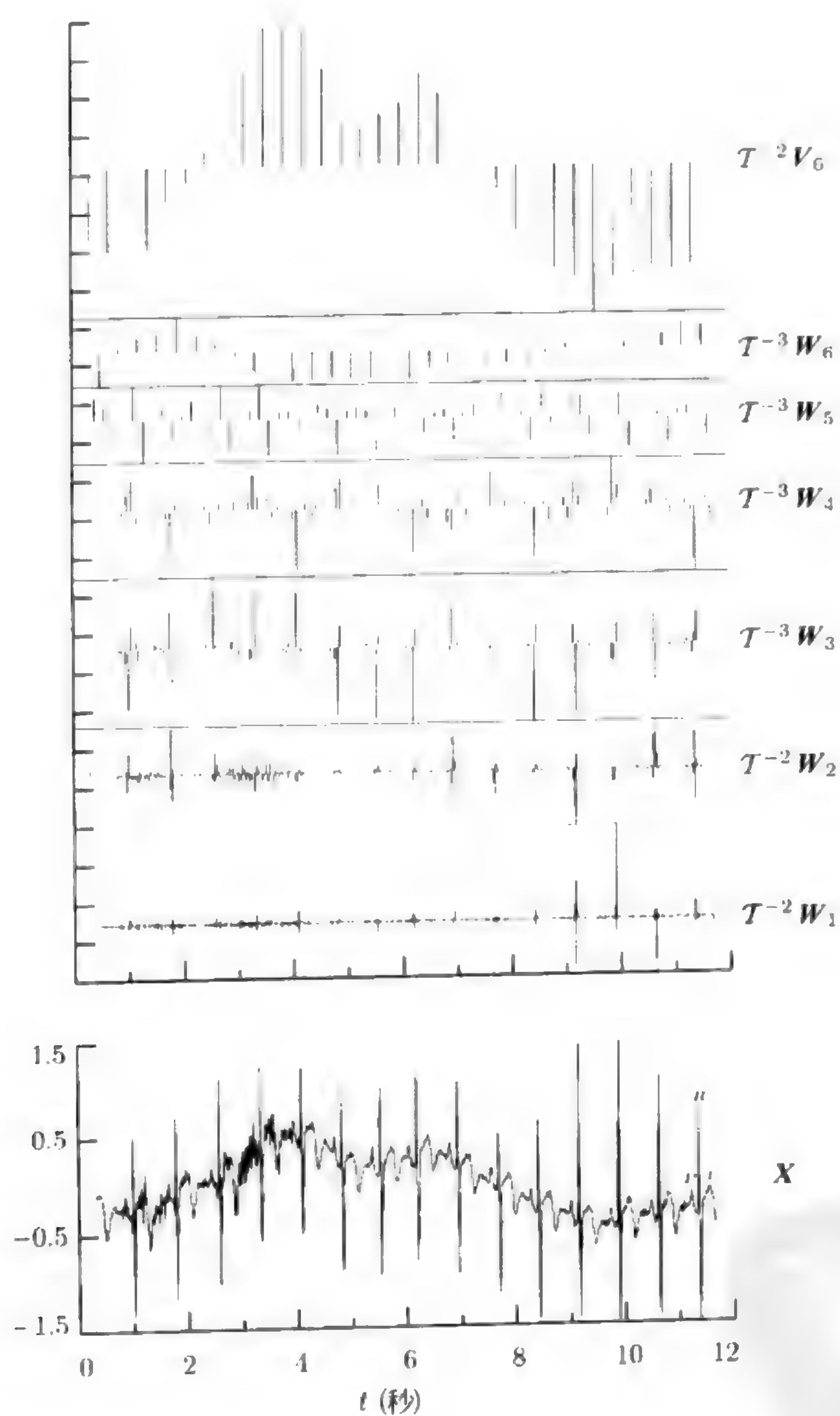


图 127　心电图时间序列(数据来自 Gust Bardy 和 Per Reinhall，华盛顿大学)的 LA(8)离散小波变换系数

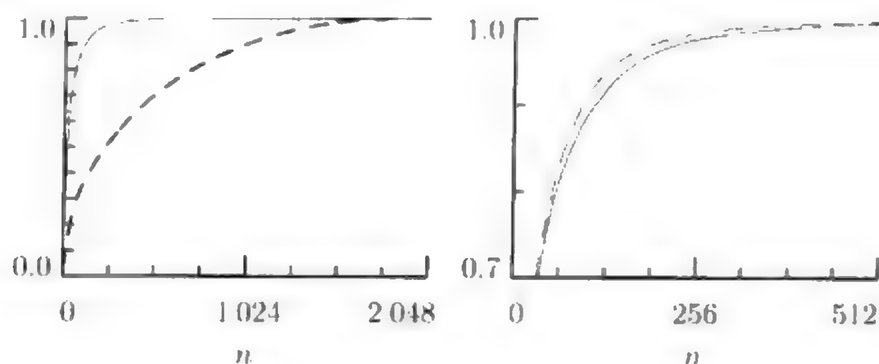


图 128 规范化部分能量序列. 左图分别绘出了原始时间序列的规范化部分能量序列(粗的点划线)、规范正交离散傅里叶变换的规范化部分能量序列(细的点划线)和哈尔离散小波变换的规范化部分能量序列(实线). 右边的图表示哈尔离散小波变换的规范化部分能量序列的比值图(实线)和 D(4)的规范化部分能量序列的比值图(虚线). C(6)和 LA(8)离散小波变换的规范化部分能量序列基本上和 D(4)的规范化部分能量序列一样

尺度系数向量 $\mathbf{V}_6$ 中. 这些系数和在尺度 $\lambda_6\Delta t=64\ \Delta t=0.36$s 上的心电图序列的平均相联系，因此反映了基线偏移. 实际上我们选择 $J_0=6$ 使得偏移可以通过部分离散小波变换分离成一个单独的成分：如果 J_0 取小的话将混合部分的 P、R 和 T 波的成分，而 J_0 取大的话将导致基线偏移成分被分解，这个分解不是我们在做心电图分析时所关心的.

图 127 给出了调整使接近零相位(参见 4.8 节讨论)的 LA(8)系数子向量 $\mathbf{W}_j(j=1,\cdots,6)$ 和 $\mathbf{V}_6$. 这个平移的结果使得图上画的 $\mathbf{W}_j$ 的元素 $W_{j,t}(t=0,\cdots,N_j-1)$的顺序和序列顺序有所不同，而是由循环平移向量 $\mathcal{T}^{-\gamma_j^{(H)}}\mathbf{W}_j$ 的元素给出，$\gamma_j^{(H)}$ 的值已在图中给出(我们将在下一节讨论如何确定这个值). 例如，当 $j=5$ 时，$N_5=64$，我们有 $\gamma_5^{(H)}=3$，这意味着我们是从左到右绘 $W_{5,3}$，$W_{5,4}$，…，$W_{5,63}$，$W_{5,0}$，$W_{5,1}$ 和 $W_{5,2}$. 这些系数和零的偏差在纵轴上标记出来了(因此 $\mathbf{W}_1$ 系数是从 -1.5 到大约 2.5；$\mathbf{W}_2$ 系数是从大约 -1.5 到大约 1.5；依次类推). 注意 X 轴上 $\mathbf{W}_2$ 和 $\mathbf{W}_3$(这些系数与物理尺度 $\tau_2\ \Delta t=2\Delta t\doteq0.011$ 和 $\tau_3\ \Delta t=4\Delta t\doteq0.022$ 之间的变化相应)的 R 波事件主要是和长钉形相联系——但不完全是. 水平的时间间隔随着尺度增大而增大，对第 j 层小波系数时间间隔为 $2\tau_j\Delta t=2^j\Delta t=2^j/180$s. 对第 j 层尺度系数时间间隔为 $\lambda_{J_0}\Delta t=2^{J_0}\Delta t=0.36$s. 粗略地讲，某个系数的位置加上或减去时间间隔的一半意味着用心电图序列以某种方式相加的结果(不论是小波系数对应的对平均做差分还是尺度系数对应的某一尺度的平均). 注意由于尺度系数和在尺度 0.36s 上的平均成正比，它们的形状大致和基线偏移差不多.

为了研究不同的离散小波变换对时间序列的刻画，图 128 给出了原始时间序列和小波变换后的时间序列(为了比较还给出了规范正交离散傅里叶变换变换后的序列)规范后的部分能量序列. 对于一个实值或复值的序列$\{U_t: t=0, \cdots, M-1\}$，一个规范化部分能量序列(NPES)是通过下述方式形成的. 首先，形成序列的平方$|U_t|^2$，然后将这些量按大小排序

$$|U_{(0)}|^2 \geqslant |U_{(1)}|^2 \geqslant \cdots \geqslant |U_{(M-2)}|^2 \geqslant |U_{(M-1)}|^2,$$

其中$|U_{(0)}|^2$表示其中M个值中最大的；$|U_{(1)}|^2$是第二大的值；如此下去直到最小的观察值$|U_{(M-1)}|^2$. 规范化部分能量序列定义为

$$C_n \equiv \frac{\sum_{u=0}^{n} |U_{(u)}|^2}{\sum_{u=0}^{M-1} |U_{(u)}|^2}, \quad n = 0, 1, \cdots, M-1$$

(注意分母等于原始的序列$\{U_t\}$的平方的和). 由定义，规范化部分能量序列是单调非降的序列，且$0<C_n\leqslant 1$对所有的n成立，$C_{M-1}=1$. 如果一个特殊的规范正交变换可以抓住时间序列的关键特征，并用一些系数表示出来，我们希望C_n对于很小的n会接近于 1. 从图 128 左边的图可以看出，因为抓住了心电图序列的特征，哈尔小波变换可以将这个序列中的能量集中起来. 用哈尔小波变换表示比序列本身或规范正交离散傅里叶变换的结果能量都更集中. 从右面的图可以看出 D(4)离散小波变换比哈尔离散小波变换能更好地集中能量(相应于 C(6)和 LA(8)的规范化部分能量序列这里没有给出，不过它们的结果基本上和 D(4)的规范化部分能量序列差不多). 显然，序列 $\mathbf{X}$ 自身的规范化部分能量序列增长非常缓慢，主要是由于基线偏移的影响. 规范正交离散傅里叶变换的规范化部分能量序列比序列 $\mathbf{X}$ 自身的规范化部分能量序列要好一些，因为它用很少的几个系数就能抓住总体的基线偏移特征，但是它不能很好地表示局部化的 P、R 和 T 波；部分离散小波变换可以同时很好地刻划基线偏移特征和局部特征 P、R 及 T 波，其中基线偏移特征用尺度系数表示，局部特征 P、R 和 T 波用小波系数表示.

图 130、131、132 和 133 分别给出了哈尔、D(4)、C(6)和 LA(8)的$J_0=6$层的多分辨分析. 细节$\mathcal{D}_j$和光滑$\mathcal{S}_6$用类似图 127 的方式给出，在这 4 幅图中其纵轴上相邻的标记都为单位长度(然而，与小波系数不同，我们没有进行相位矫正，即对$\mathcal{D}_j$和$\mathcal{S}_6$进行前移). 对于这四个滤波器，每个细节$\mathcal{D}_j$都有零均值，而每个光滑$\mathcal{S}_6$的均值都和原始序列 $\mathbf{X}$ 的样本均值相同. 每个小波相应的多分辨分析满足如下附加条件：

$$\mathbf{X} = \sum_{j=1}^{6} \mathcal{D}_j + \mathcal{S}_6.$$

直观地看，LA(8)的多分辨分析从视觉上看比其他的三个小波都要好：哈尔的高层细节和光滑有呈现出明显的块状，而 D(4)和 C(6)小波有一些“鲨鱼鳍”和“三角形”的痕迹，在细节$\mathcal{D}_6$和光滑$\mathcal{S}_6$中表现得很明显. 为了理解为什么这些人

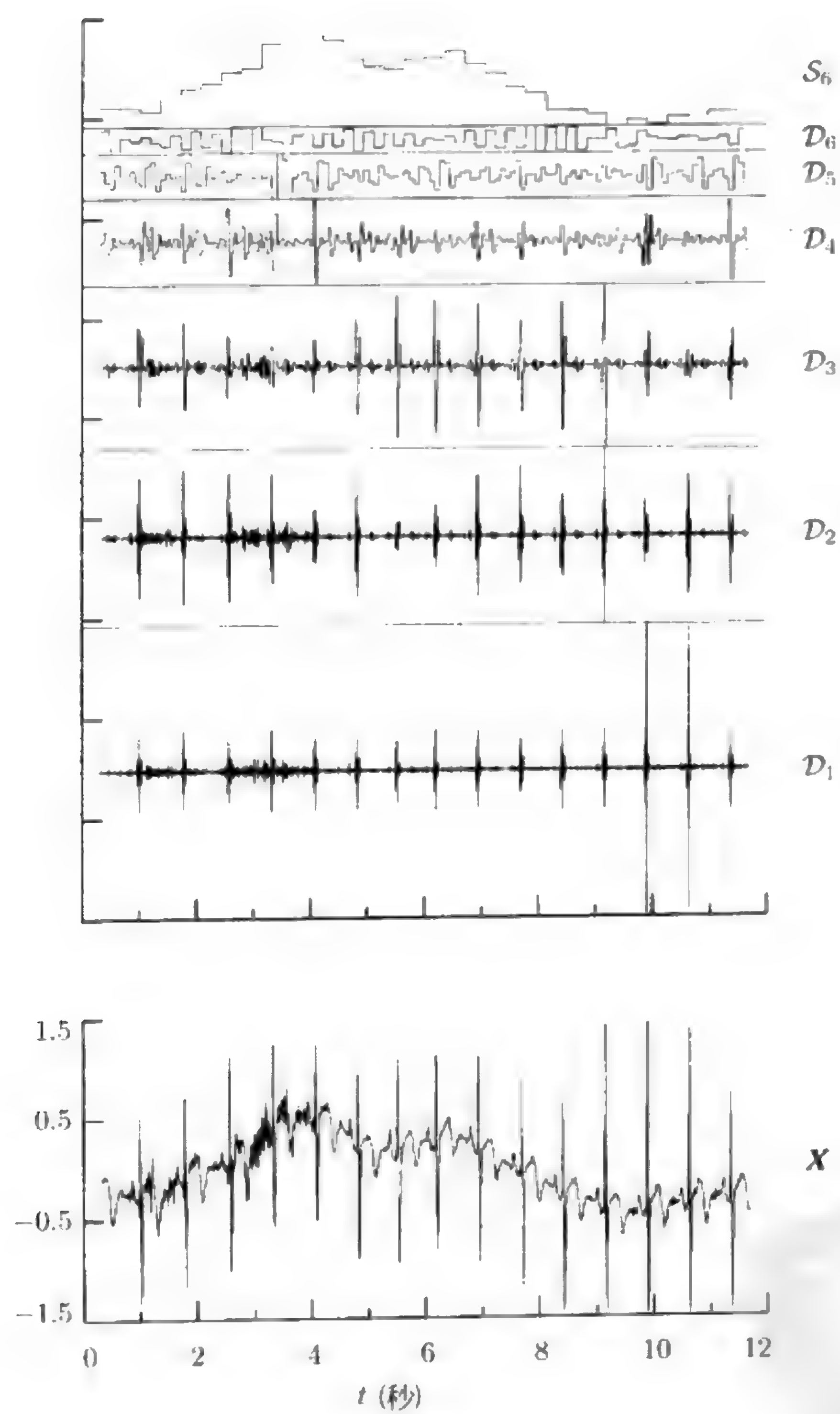


图 130 心电图时间序列的哈尔离散小波变换的多分辨分析(详情参见正文)

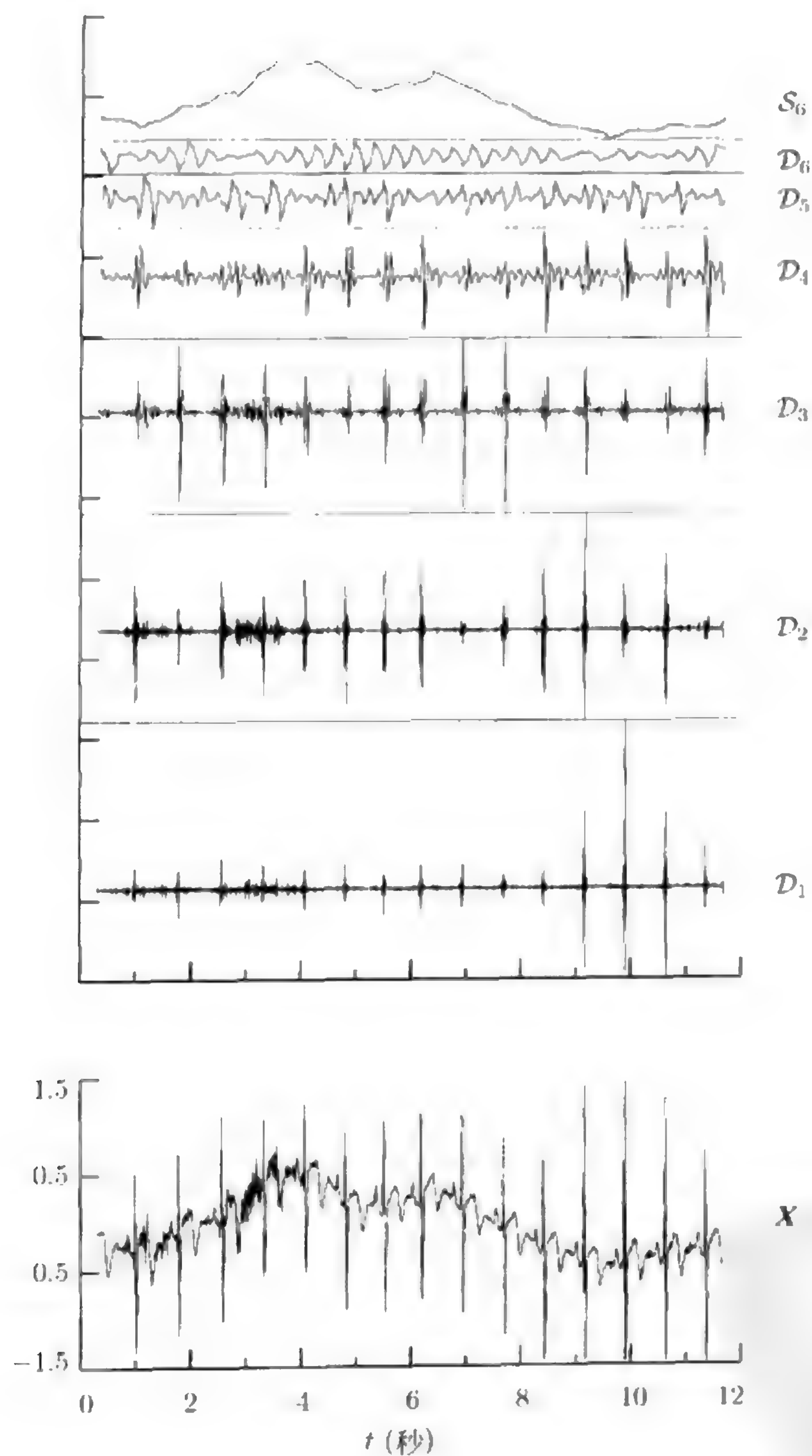


图 131　心电图时间序列的 D(4)离散小波变换的多分辨分析(详情参见正文)

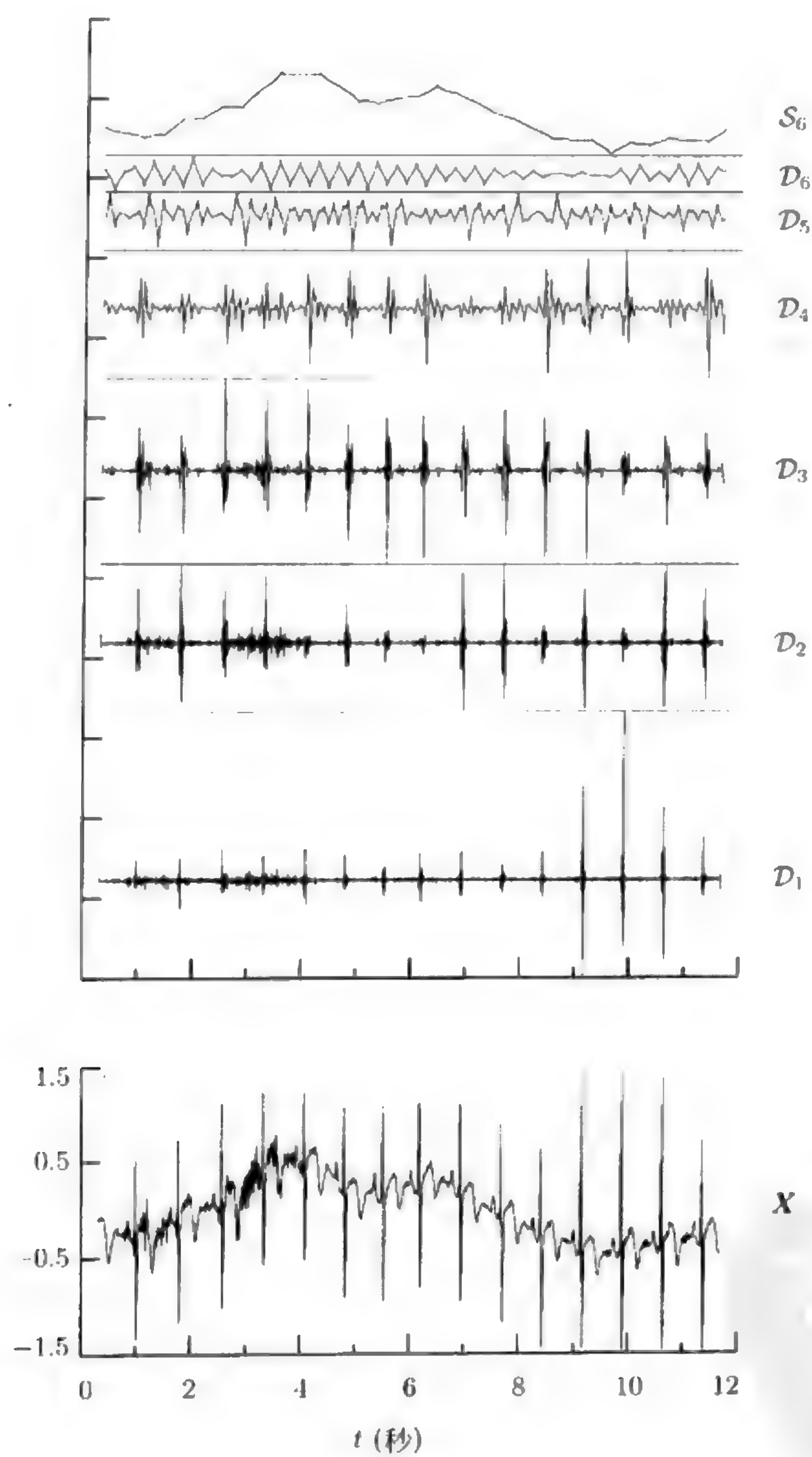


图 132 心电图时间序列的 C(6)离散小波变换的多分辨分析(详情参见正文)

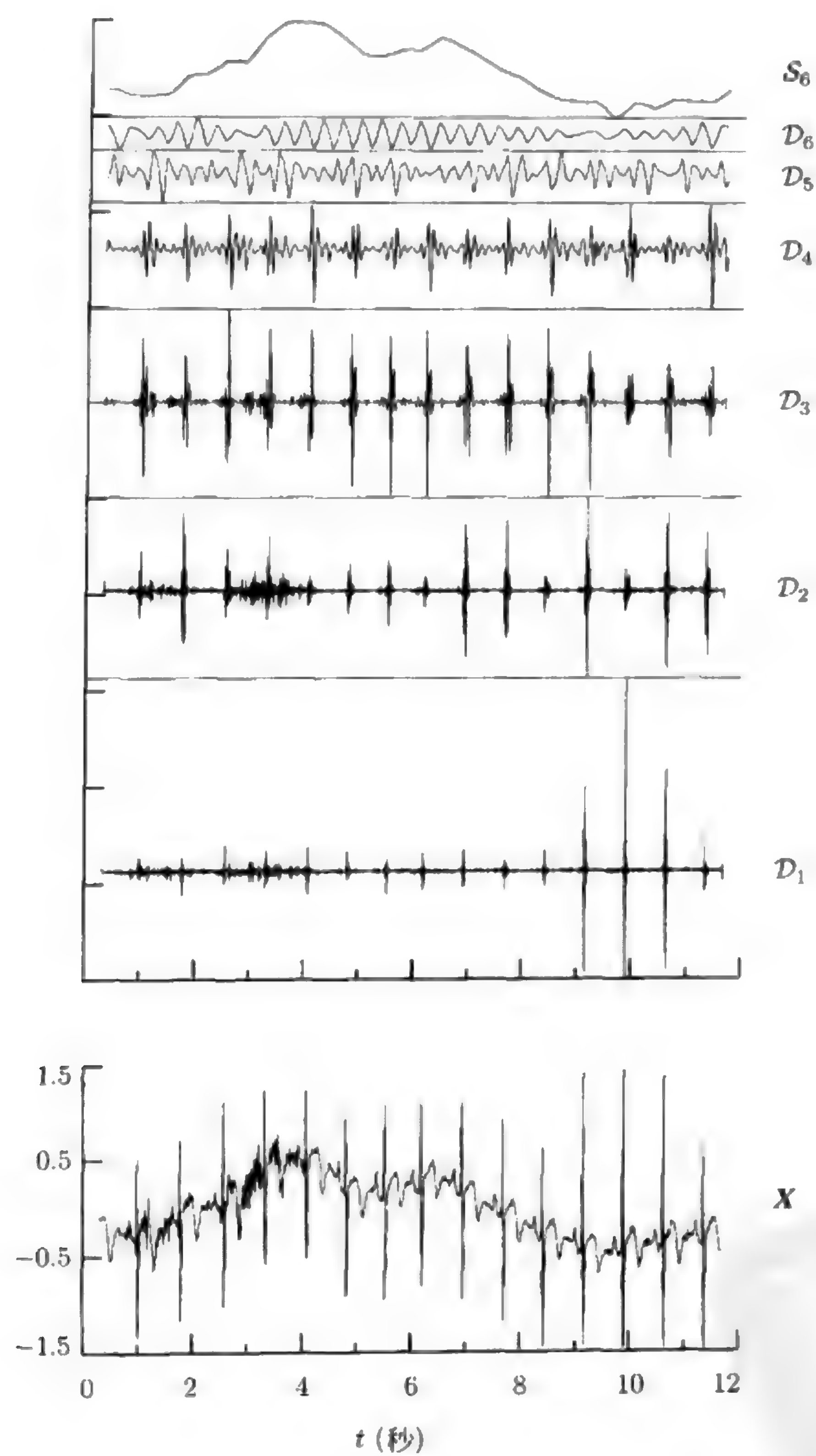


图 133　心电图时间序列的 LA(8) 离散小波变换的多分辨分析(详情参见正文)

工痕迹会出现，回忆每个细节能表示为

$$\mathcal{D}_j = \mathcal{W}_j^T \mathbf{W}_j,$$

因此$\mathcal{D}_j$是一个在尺度$\tau_j \Delta t$上的$\mathcal{W}$各行向量转置的线性组合．这些行向量包含了第j层小波滤波器$\{h_{j,l}\}$，对于 D(4)和 C(6)小波分别在图 98a 和图 125 中给出了其图形．高层 D(4)小波滤波器的“鲨鱼鳍”特征和高层 C(6)小波滤波器的“三角形”特征因此会随着多分辨分析中比较大的$W_{j,t}$而表现出来(类似地，哈尔小波的细节和光滑由于哈尔小波滤波器的特征而表现出块状)．而光滑且接近对称的 LA(8)在高层小波滤波器(参见图 98b)看起来更匹配心电图数据，它不会产生明显受小波形状影响的多分辨分析成分，因此这样的多分辨分析成分更能表现原始数据本身的特征.

基于尺度特性的多分辨分析可以有效地将心电图数据分解成各种成分以利于进一步分析．例如，由于$\mathcal{D}_5$和$\mathcal{D}_6$反映了在物理尺度$\tau_5 \Delta t = 16\Delta t \doteq 0.09$s 和$\tau_6 \Delta t = 32\Delta t \doteq 0.18$s 上的变化，这些细节提取了 P 和 T 波的成分．注意 LA(8)的光滑$\mathcal{S}_6$很好地反映了基线偏移．最后，由病人的运动导致的短尺度的瞬间波动主要集中在细节$\mathcal{D}_1$和$\mathcal{D}_2$中.

总的来说，$J_0 = 6$层的 LA(8)的部分离散小波变换比哈尔、D(4)和 C(6)得到了一个更好的多分辨分析．和序列 $\mathbf{X}$ 本身和它的规范正交离散傅里叶变换比较起来，所有的离散小波变换都可以更紧凑地表示心电图数据，这种紧凑程度用规范化部分能量序列来衡量．以这种标准看 D(4)、C(6)和 LA(8)离散小波变换都要比哈尔离散小波变换好．其实对于 LA(8)多分辨分析，和“极大重叠”离散小波变换相比较也存在一些明显的人工痕迹．“极大重叠”离散小波变换是第 5 章内容(在 5.7 节我们还会用到心电图的例子).

4.10 节的评论与扩展

[1]我们应该注意到心电图时间序列是通过有限精度的仪器来采集心脏跳动的．第一，在图 127 的下图给出的图形说明了有 15 个 R 波(尖峰)，其中在 9 和 11 秒之间的三个 R 波非常大．仔细看这三个 R 波我们可以发现，实际上这三个 R 波其实都有两个峰，在图 127 中表现为一个波因为水平坐标的分划不够细(另外 12 个 R 波都是单峰的)．在这三个 R 波中，实际上真正的 R 波是两个峰中的第一个峰，第二个峰被认为是仪器的一个缺陷，因为这两个峰太接近了，间隔太短，不能认为是心脏跳动引起的波峰．第二个峰包含了两个相似的突出，第一个突出具有偶数指标，第二个突出具有奇数指标，这使得这对突出不能够被哈尔离散小波变换系数$\mathbf{W}_1$检测出来．这些突出如果具有指标奇/偶的形式，而不是偶/奇

的形式，就可以被检测出来. 在第 5 章，我们将考虑对离散小波变换做一点变化，其中一个原因是这样做可以消除对齐效应.

第二，在每个 R 波出现以后，心电图数据迅速衰减变为一个很窄的谷(所谓的"QRS 复合体"的"S 部分"，其中 Q 和 S 是在 R 波、P 波和 T 波间的谷，这些波在图 127 下图中标记出来了). 出于某些技术上的原因，仪器采集的数据需要通过一个非线性滤波器，这个非线性滤波器对两个相邻的时间点上采集的数据的变换率会有限制. 这个滤波器使得我们不可能精确地采集迅速衰减的数据. 这个缺陷在图 126 的哈尔离散小波变换的 $\mathbf{W}_1$ 上有所体现，其中有 14 个幅度几乎相同的谷(图 130 中的多分辨分析的 $\mathcal{D}_1$ 成分有几乎类似的缺陷)，因此我们论证了哈尔小波不是很适合分析心电图数据，但是用它可以分析出仪器的这个细小的缺点.

4.11 实际应用中需要考虑的问题

就像上一节例子中所说的那样，为了在实际中有效地应用小波变换分析时间序列，我们需要考虑很多实际因素. 这里将讨论如何做一些选择，如何处理边界问题和长度不是 2 的整数次幂的情形(见 Bruce and Gao，1996a，还有 Ogden，1997，也讨论了一些有关实际因素的考虑).

• 小波滤波器的选择

我们面对的第一个问题是如何从 4.8 节和 4.9 节讨论的众多小波滤波器中选择一个特定的滤波器来进行时间序列的小波分析. 我们从前面的心电图的例子已经看出，小波滤波器的合理选择依赖于特定的应用，所以后面几章还要给出小波分析时间序列的例子，因此在这里我们将重新考虑选择小波这个问题. 那几章我们选择的小波说明了特定的分析目标(例如分离时间序列中的瞬间变化、信号估计、长记忆过程的参数估计、检验方差的一致性、估计小波方差等等)，及需要在一个小波滤波器中达到那种目的的性质.

研究后面几章的例子可以学习如何选择小波滤波器(特别是，参见 5.8、5.9、8.6、8.9、9.7 和 9.8 节)，我们在这里只做一个整体的描述. 一般地说，我们的选择主要是要平衡两种考虑. 一方面，宽度很短的小波滤波器(如 $L=2, 4, 6$)有时会在分析结果中引入我们不希望的人工痕迹(在心电图的例子中，哈尔、D(4)和 C(6)多分辨分析导致了不真实的块状、鲨鱼鳍状和三角形状(参见图 130、131 和 132)，但是对于某些应用来说哈尔小波也很合适(参见 8.6 节和 9.6 节的两个例子). 另一方面，小波滤波器的 L 如果大的话可以更好地反映时间序列的特征，但是使用这些滤波器可以导致：(i)更多的系数被边界所影响，(ii)离散小波变换系数的局部化等级有些减少，(iii)计算量增大. 一个合理的方案是要选最小的可以得到满意结果的 L. 在实际中，选择的

表 136

L	L'_1	L'_2	L'_3	L'_4	$L'_{j\geq 5}$
2	0	0	0	0	0
4	1	2	2	2	2
6	2	3	4	4	4
8	3	5	6	6	6
10	4	6	7	8	8
12	5	8	9	10	10
14	6	9	11	12	12
16	7	11	13	14	14
18	8	12	14	15	16
20	9	14	16	17	18

注：宽为 L 的小波滤波器滤波后得到的 $\mathbf{W}_j$ 和 $\mathbf{V}_j$ 的边界系数的个数 L'_j（这里我们假定 $L'_j \leq N_j$，其中 $N_j = N/2^j$ 是 $\mathbf{W}_j$ 和 $\mathbf{V}_j$ 的长度）. 边界系数就是多少都受到边界条件影响的系数.

过程就是比较初步的分析，不断增大 L，直到我们得到一个小波系数没有人工痕迹的分析结果. 如果我们还需要得到的离散小波变换系数在时间上是对齐的(即平移相位使接近零相位)，这个选择方案总是导致最后的最佳选择为 LA(8)小波滤波器(如果小波系数的对齐是很重要的，那么最好使用最接近对称小波或 coiflet 小波，如果选择了最接近对称小波，那么最好选择一个长度为 4 的倍数的最接近对称小波——这样的小波滤波器和尺度滤波器的相位更接近线性相位，参见图 115).

• 处理边界条件

一个部分离散小波变换或完全离散小波变换需要使用循环滤波，因此滤波运算接近序列 $\mathbf{X}$ 的起始部分和最后的部分("边界")被看成是一个周期为 N 的序列的一部分. 在循环滤波中，我们实际上假设了 X_{N-1}，X_{N-2}，…等于 X_{-1}，X_{-2}，…，这个假设对有些序列来说如果序列的长度选的合适的话是合理的. 例如，5.8 节讨论的海洋潮水水位序列就有很强的周期性，如果序列的时间跨度近似为一年的话这个序列就可以当成周期序列处理. 对于其他一些序列，特别是对在 X_{N-1} 和 X_0 两点值变化很大的序列来说，周期性假设是有问题的.

因为周期性是一个有问题的假设，我们需要仔细考虑这种假设会如何影响时间序列的小波分析结果. 为了做到这一点，我们首先对周期性假设对离散小波变换系数和相应的多分辨分析产生的影响做一个量化. 对于小波系数，我们在本节的评论与扩展的[1]中给出了 N_j 维向量 $\mathbf{W}_j$ 的边界系数的个数（即被序列循环以后受到影响的系数)为 $\min\{L'_j,\ N_j\}$，其中 L'_j由式(146a)定义，在表 136 中给出了 $L=2$，4，…，20，(对于尺度系数也有类似的结果). 从这个表中很明显可以看出，L'_j不依赖于 N，且随着 L 的增大而增大，随着 j 的增大 L'_j也是非降的，且满

表 137a

L	$\bar{\gamma}_1^{(H)}$，$\gamma_1^{(H)}$	$\bar{\gamma}_2^{(H)}$，$\gamma_2^{(H)}$	$\bar{\gamma}_3^{(H)}$，$\gamma_3^{(H)}$	$\bar{\gamma}_4^{(H)}$，$\gamma_4^{(H)}$	$\bar{\gamma}_{j\geq 5}^{(H)}$，$\gamma_{j\geq 5}^{(H)}$
8	1，2	3，2	3，3	3，3	3，3
10	2，2	3，3	4，3	4，4	4，4
12	2，3	4，4	5，4	5，5	5，5
14	2，4	4，5	6，5	6，6	6，6
16	3，4	6，5	7，6	7，7	7，7
18	4，4	6，6	7，7	8，7	8，8
20	4，5	7，7	8，8	9，8	9，9

注：在 $\mathcal{T}^{-\gamma_j^{(H)}}\mathbf{W}_j$ 的开头和结尾的最接近对称边界小波系数的个数(假定 $L_j'\leqslant N_j$)．开始部分的边界系数用 $\bar{\gamma}_j^{(H)}$ 来记．最后部分的边界系数用 $\gamma_j^{(H)}$ 来记．注意 $\bar{\gamma}_j^{(H)}+\gamma_j^{(H)}=L_j'$，为边界系数的总数量(参见表 136)．

表 137b

L	$\bar{\gamma}_1^{(G)}$，$\gamma_1^{(G)}$	$\bar{\gamma}_2^{(G)}$，$\gamma_2^{(G)}$	$\bar{\gamma}_3^{(G)}$，$\gamma_3^{(G)}$	$\bar{\gamma}_4^{(G)}$，$\gamma_4^{(G)}$	$\bar{\gamma}_{J_0\geq 5}^{(G)}$，$\gamma_{J_0\geq 5}^{(G)}$
8	2，1	3，2	4，2	4，2	4，2
10	2，2	3，3	3，4	4，4	4，4
12	3，2	5，3	5，4	6，4	6，4
14	4，2	6，3	7，4	8，4	8，4
16	4，3	6，5	7，6	8，6	8，6
18	4，4	6，6	7，7	7，8	8，8
20	5，4	8，6	9，7	9，8	10，8

注：像表 137a，但现在是对于最接近对称的尺度系数．同样我们有 $\bar{\gamma}_{J_0}^{(G)}+\gamma_{J_0}^{(G)}=L_{J_0}'$，其中 L_{J_0}' 由表 136 给出．

足如下不等式：

$$\frac{L}{2}-1\leqslant L_j'\leqslant L-2$$

($j=1$ 时 L_j'等于下界，j 足够大时达到上界)．当 N 增大的时候 N_j 增大，边界系数比值 L_j'/N_j 变为任意小；另一方面，对于固定的 N，这个比值随着 j 的增大而增大（除了哈尔小波($L=2$)的情形，此时离散小波变换不受周期假设影响）．此外，当 $0<L_j'\leqslant N_j$ 时，$\mathbf{W}_j$ 或 $\mathbf{V}_j$ 的边界系数是每个向量中的前 L_j'元素．

我们现在考虑最接近对称滤波器和 coiflet 滤波器(假设 $L_j'\leqslant N_j$)．如果照着 4.8 节描述的方法，并对 $\mathbf{W}_j$ 做前移(为了绘图)，我们看对于 L_j'个边界小波系数会发生什么．我们在本节评论与扩展中的[2]将给出与 $\mathcal{T}^{-\gamma_j^{(H)}}\mathbf{W}_j$ 中的元素相应的绘图顺序，其中 $0<\gamma_j^{(H)}<L_j'$．这意味着 L_j'个边界系数的循环平移量 $\gamma_j^{(H)}$ 使得它

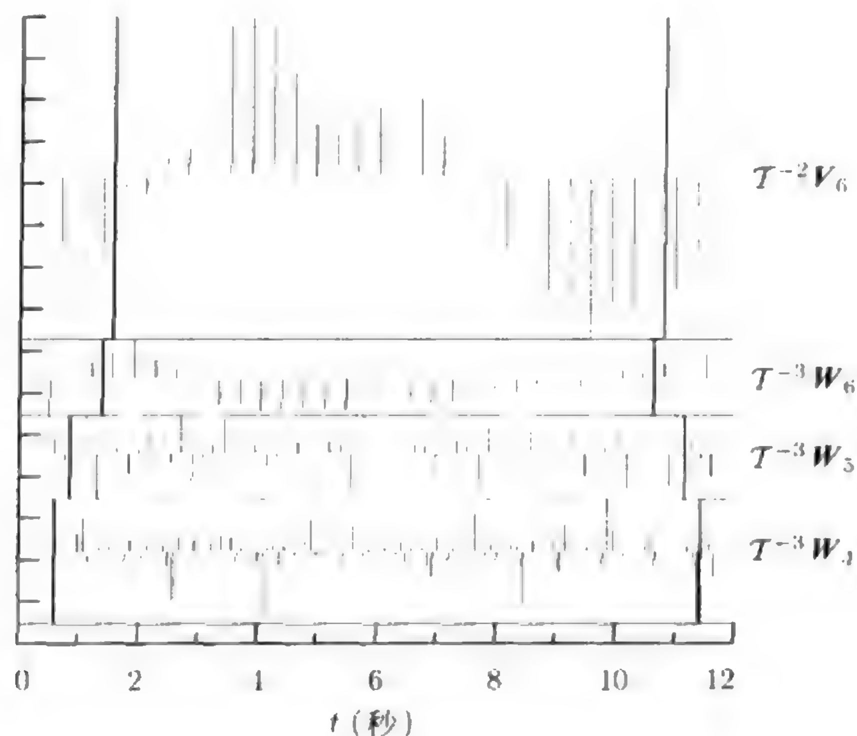


图 138 心电图时间序列的 LA(8) 离散小波变换的四个循环前移的子向量(参见图 127). 在两条粗的垂直线之间的子向量不受循环的影响，而在这两条线之外的是边界系数，绘出的边界系数的个数与表 137a 和表 137b 的第一行是相同的

们跑到了图的后面，而剩下的 $\tilde{\gamma}_j^{(H)} \equiv L'_j - \gamma_j^{(H)}$ 个边界系数在图的前面部分. 对在本书中讨论的宽为 L 的所有最接近对称滤波器，表 137a 给出了 $\tilde{\gamma}_j^{(H)}$ 和 $\gamma_j^{(H)}$ 的值(类似的对尺度系数定义的量 $\tilde{\gamma}_{J_0}^{(G)}$ 和 $\gamma_{J_0}^{(G)}$ 在表 137b 中给出). 练习[4.27]对 coiflet 滤波器给出了类似的表.

处理周期假设的一个方法是在离散小波变换系数的图上标出那个值是边界值. 例如，图 138 重新给出了图 127 中的心电图序列的 $j=4$, 5, 6 层 LA(8)小波系数(对于 $J_0=6$ 的尺度系数也给出来了). 我们已经加了粗的垂直线，在这条线的外面是边界系数. 例如，在$\mathcal{T}^{-3}\mathbf{W}_6$ 的前面部分有三个边界系数，这和表 137a 中的数据是吻合的；类似地，在 $\mathcal{T}^{-2}\mathbf{V}_6$ 的前面部分有四个边界系数，也和表 137b 中的数据吻合.

下面我们考虑周期化是如何影响细节 $\mathcal{D}_j$ 的(这里我们不再限制仅仅是最接近对称滤波器和 coiflet 滤波器了). 我们在下面的评论与扩展[3]中将给出，除了由哈尔小波产生的$\mathcal{D}_j$，其他的情况 $\mathcal{D}_j$ 的系数的前面部分和后面部分都有受循环影

响的系数. 这些系数的指标为

$$t = 0,\cdots,2^j L'_j - 1 \quad 及 \quad t = N-(L_j-2^j),\cdots,N-1; \tag{139}$$

即$\mathcal{D}_j$ 的前 $2^j L'_j$个元素和后 $L_j - 2^j$ 个元素受到了循环的影响(类似的结果对$\mathcal{S}_{J_0}$也成立). 于是我们看到$\mathcal{D}_j$ 中的边界元素依赖于 L 和 j 但不依赖于 N, 随着 L 或 j 的增加, 更多的元素被影响了. 当 $j=1$ 时, $L'_1=\frac{L}{2}-1$ 且 $L_1=L$, 下界和上界就成了

$$t = 0,\cdots,(L-2)-1 \quad 及 \quad t = N-(L-2),\cdots,N-1,$$

因此在$\mathcal{D}_1$ 的两端有 $L-2$ 个元素受到了循环的影响. 当 j 足够大时, L'_j达到极大值$L-2$, 我们有 $2^j L'_j = L_j - 2^j + L - 2$, 因此两端将有 $L-2$ 个或更多的边界元素.

作为一个例子, 我们考虑式(81c)中的$\mathcal{D}_1$. 这时 $j=1$, $L=4$. 我们上面看到$W_{1,0}$是一个边界系数, 因此$\mathcal{D}_1$ 的前面的两个元素将会被矩阵$\mathcal{W}_1^{\mathrm{T}}$ 所影响. 这和式(139)是相容的, 因为我们能用表 136 得到 $2^j L'_j = 2L'_1 = 2$. 此外, $\mathcal{W}_1^{\mathrm{T}}$ 的最后两行包含了周期化, 因此$\mathcal{D}_1$ 的最后两个元素也被影响了. 这也和式 139 式相吻合, 因为 $L_j - 2^j = L - 2 = 2$.

因为与离散小波变换的系数相联系, 所以我们用垂直线标定$\mathcal{D}_j$ 与$\mathcal{S}_{J_0}$的边界范围. 这在图 140 中标出了(是对图 133 的重绘). 对比这个图的垂直标记和图 138, 我们可以看到细节$\mathcal{D}_j$ 受影响的元素比 $\mathbf{W}_j$ 受影响的元素要多, 下面的练习将这一结果量化.

练习[139] 证明, 对于大的 j 和 N, 在$\mathcal{D}_j$ 中受影响的元素和 $\mathbf{W}_j$ 中受影响的元素的比例接近于 2. ◁

类似的结论对$\mathcal{S}_{J_0}$和 $\mathbf{V}_{J_0}$成立.

应当强调的是, 虽然我们知道 $\mathbf{W}_j$、$\mathbf{V}_{J_0}$、$\mathcal{D}_j$ 和$\mathcal{S}_{J_0}$中哪些元素受到了循环的影响, 事实上循环导致的影响可以非常小, 特别是当第一个系数和最后一个系数相差不大时. 作为一个例子, 在图 140 中描绘的多分辨分析中, 光滑$\mathcal{S}_6$ 描绘的是心电图序列在尺度 $2^6\Delta t = 0.36$s 上的平均. 比较这个序列和$\mathcal{S}_6$ 在垂直线标记前后的形状, 我们看到$\mathcal{S}_6$ 和心电图序列非常吻合, 视觉上看是大约三分之一秒的尺度上的平均. 循环产生的影响在边界区域非常小. 我们应当意识到垂直线标记的位置很大程度上取决于$\{g_{j,l}\}$的宽度 L_j, 但是事实上这个等价滤波器有很多很小的元素(参见图 98b). 除非 X_0 和 X_{N-1}有很大的差别, 否则实际上受到循环影响比较大的部分比垂直线标记的受循环影响的区域要小的多. 因此垂直线标记的边界区域是对受循环影响的区域的一个保守度量.

现在我们看如何来决定小波分析结果受边界影响的部分, 让我们现在考虑减小循环影响的方法. 因为离散傅里叶变换也是将时间序列看成是周期的, 我们可

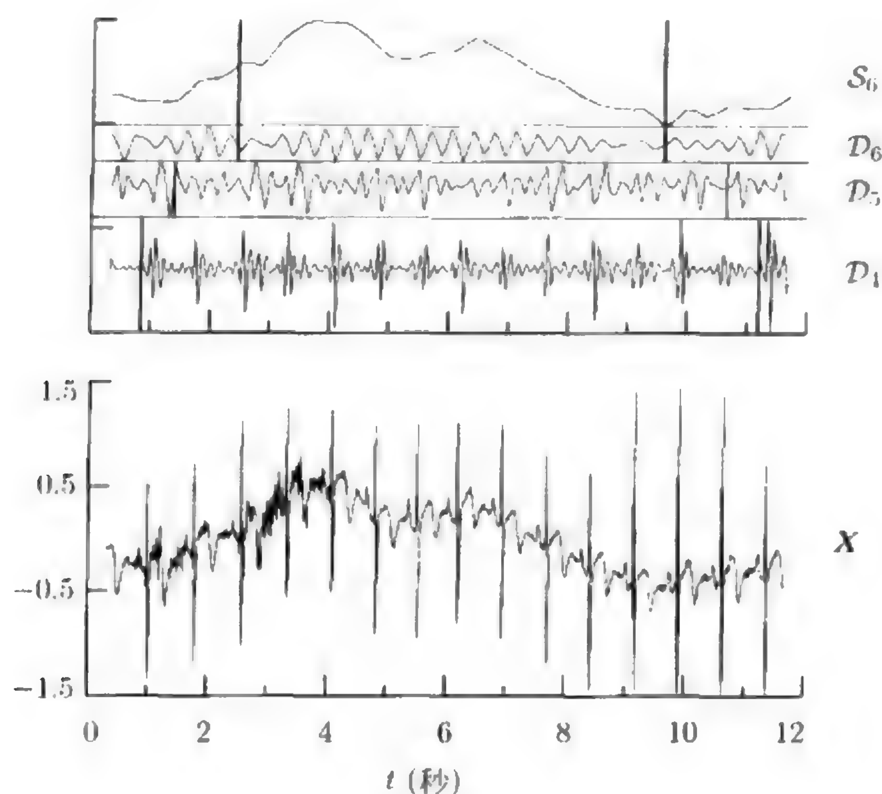


图 140　心电图时间序列的 LA(8) 离散小波变换多分辨分析的一部分(完全分析在图 133 中给出). 粗的垂直线标记出了细节 $\mathcal{D}_j$ 和光滑 $\mathcal{S}_6$ 的边界区域(即受到循环影响的部分)

以借用傅里叶分析中的一些应用成功的技巧. 其中一个技巧——非常有用——是用长为 $2N$ 的序列代替 $\boldsymbol{X}$ 作为分析目标，这个序列是通过将 $\boldsymbol{X}$ 和一个反序的 $\boldsymbol{X}$ 连接起来得到的，即

$$X_0, X_1, \cdots, X_{N-2}, X_{N-1}, X_{N-1}, X_{N-2}, \cdots, X_1, X_0. \tag{140}$$

注意这个长为 $2N$ 的序列和 $\boldsymbol{X}$ 具有相同的样本均值. 对 $\boldsymbol{X}$ 进行循环滤波相当于用 X_{N-1}，X_{N-2}，…代替 X_{-1}，X_{-2}，…；对上式中的序列进行循环滤波相当于用 X_0，X_1，…代替 X_{-1}，X_{-2}，…，这样就消除了由于 X_0 和 X_{N-1}之间的不连续导致的边界效应. 当直接对 $\boldsymbol{X}$ 进行离散小波变换时，我们称为用循环边界条件对 $\boldsymbol{X}$ 进行分析，当对 $\boldsymbol{X}$ 对称延拓后再进行离散小波变换时，我们称为用对称延拓条件(反射条件)对 $\boldsymbol{X}$ 进行分析. 用对称延拓条件分析序列的时候我们付出的代价是计算时间的延长和存储量的增大，不过这些通常不会有大的问题.

作为例子，图 142 给出了 LA(8)对心电图序列用对称延拓条件进行分析后得到的多分辨分析. 这里我们计算了基于心电图序列的长为 $N=4\ 096$ 的序列的细节和光滑，延拓过程如式(140). 我们将得到的细节和光滑截断一半得到图 142，使得该图和图 133(用循环边界条件计算的结果)可以做对比. 实际上这两幅图除

了前面部分和最后部分外没有什么差别，其中用对称延拓边界条件可以减小由 X_0 和 X_{N-1} 的不连续带来的一些边界效应(比较图 142 和图 140 的顶部可以看出，除了非常特殊的情形之外，在边界区域没有特别明显的不同，这再次强调了我们以前说明的一个事实，那就是循环造成的明显的影响区域比实际的边界区域小的多).

有几种方法可以处理循环带来的问题，包括在序列的两边进行多项式外推，设计特殊的"边界小波"，还有就是对对称延拓方案做一些修改：将序列先做差分，然后像式(140)那样对称延拓，接着累加起来(对于近似线性的序列来说做一次差分是有用的，对于近似二次曲线的序列来说可以做二次差分，依次类推).如果想深入了解这些处理方法，参见 Bruce and Gao(1996a)的第 4 章，Taswell and McGill(1994)，Cohen 等(1993)，还有 Greenhall 等(1999).

• 处理长度不是 2 的整数幂的序列

前面我们描述的"完全"离散小波变换要求处理的序列长度是 2 的整数幂，即 $N=2^J$，J 为某个正整数.我们在 4.7 节注意到用 J_0 层部分离散小波变换的限制稍微减弱了一些，只需要序列的长度为 2^{J_0} 的整数倍.当 N 不是 2^{J_0} 的整数倍时，我们可以用下面将要介绍的三种方法之一来得到对任意整数都适用的一种类似离散小波变换的变换.这些方法对很多特殊的应用都是适用的.注意离散小波变换的一个合理的替代是第 5 章将要介绍的极大重叠离散小波变换(MODWT)，它对任意的 N 都适用.虽然极大重叠离散小波变换不是一个规范正交变换，但是它的性质和规范正交变换很相似(比如可以导出精确的方差分析(ANOVA)和多分辨分析)，在某些方面它表现得比离散小波变换要好(比如它的多分辨分析对应零相位滤波器，保证了 $\mathbf{X}$ 和它的多分辨分析之间的对齐，这对基于多分辨分析的离散小波变换来说是不一定成立的).

假定我们现在想要对一个长为 N 的序列计算 J_0 层部分离散小波变换，而 N 不是 2^{J_0} 的倍数.令 N' 是比 N 大且为 2^{J_0} 的倍数的最小整数.第一种方法就是用某种方式延拓 $\{X_t\}$，得到新的序列 $\{X'_t\}$，其长度为 $N'>N$，然后对延拓的序列做离散小波变换(这个方法在傅里叶分析中很常用，将序列延拓的成 2 的整数幂长在快速傅里叶变换中经常出现).一种很自然的延拓方式是将延拓部分都设成序列 $\{X_t\}$ 的均值，即

$$X'_t \equiv \begin{cases} X_t, & t=0,\cdots,N-1; \\ \overline{X}, & t=N,\cdots,N'-1. \end{cases}$$

因为 $\{X'_t\}$ 的样本均值也等于 $\overline{X}$，样本方差 $\hat{\sigma}^2_X$ 和 $\hat{\sigma}^2_{X'}$ 之间的关系非常简单：

$$\hat{\sigma}^2_{X'} \equiv \frac{1}{N'}\sum_{t=0}^{N'-1}(X'_t-\overline{X})^2 = \frac{1}{N'}\sum_{t=0}^{N-1}(X_t-\overline{X})^2 = \frac{N}{N'}\hat{\sigma}^2_X.$$

如果令 $\mathbf{W}'_j$ 和 $\mathbf{V}'_{J_0}$ 表示序列 $\{X'_t\}$ 的离散小波变换子向量，则可以利用上式和

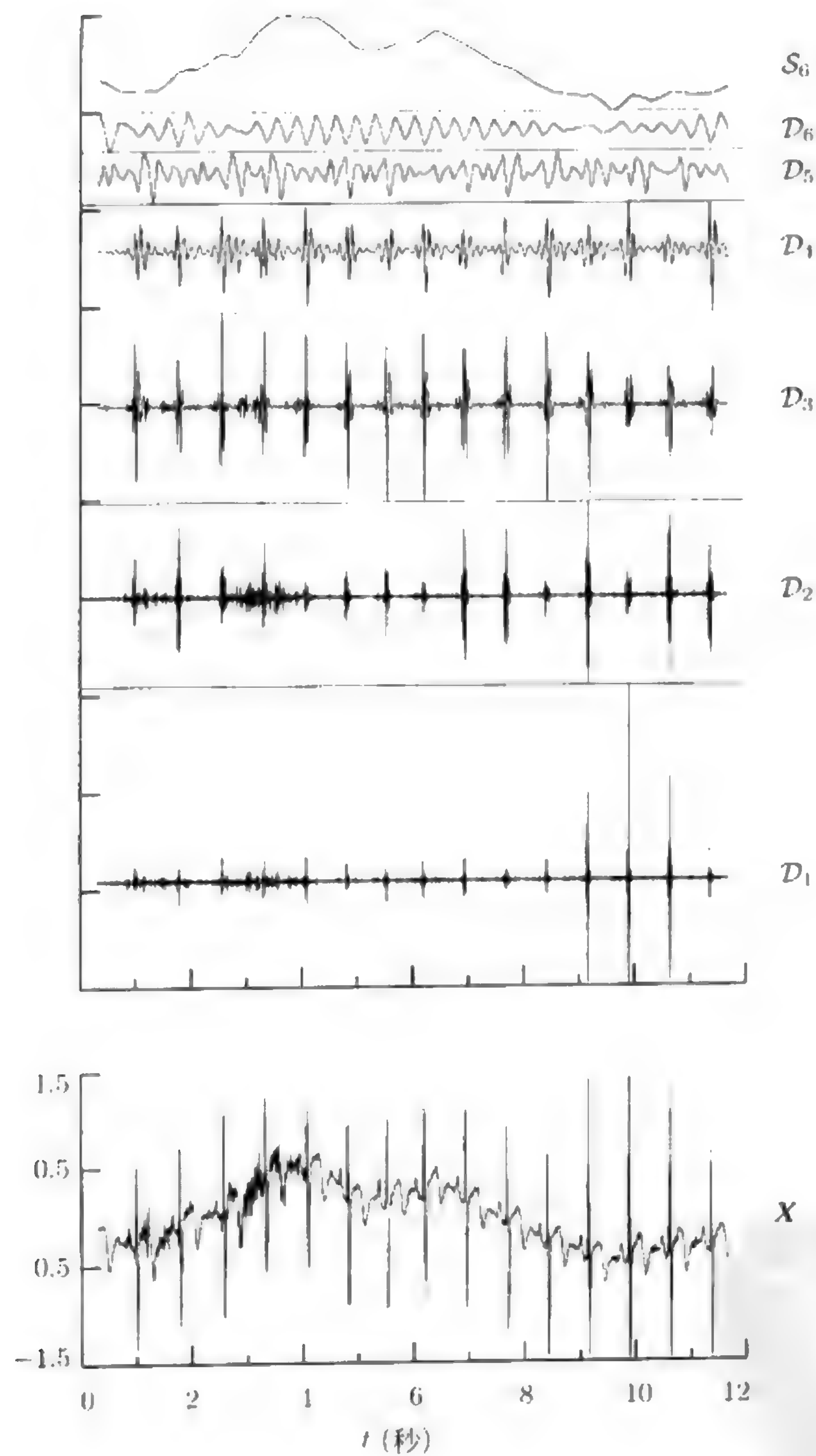


图 142 使用对称延拓边界条件的心电图时间序列的 LA(8) 离散小波变换多分辨分析（参见图 133）

式(104c)得到基于$\{X'_t\}$的方差分析的$\{X_t\}$的方差分析：

$$\hat{\sigma}_X^2 = \frac{1}{N}\sum_{j=1}^{J_0} \| \mathbf{W}'_j \|^2 + \frac{1}{N}\| \mathbf{V}'_{J_0} \|^2 - \frac{N'}{N}\overline{X}^2.$$

在上面我们将$\frac{1}{N}\| \mathbf{W}'_j \|^2$解释为样本方差$\hat{\sigma}_X^2$在尺度$\tau_j$上的变化上的分量．令$\mathbf{X}'$是元素为$\{X'_t\}$的向量，我们可以用$\mathbf{W}'_j$和$\mathbf{V}'_{J_0}$以通常的方式来形成多分辨分析：

$$\mathbf{X}' = \sum_{j=1}^{J_0} \mathcal{D}'_j + \mathcal{S}'_{J_0}.$$

因为这个附加的分解基于每一点都是有效的，通过$\mathcal{D}'_j$和$\mathcal{S}'_{J_0}$的前 N 个元素可以得到 $\mathbf{X}$ 的多分辨分析：

$$\mathbf{X} = \sum_{j=1}^{J_0} I_{N,N'}\mathcal{D}'_j + I_{N,N'}\mathcal{S}'_{J_0}, \tag{143}$$

其中$I_{N,N'}$是$N\times N'$阶矩阵，当$i=j$时，其(i,j)元素为1，当$i\neq j$时，其(i,j)元素为零.（也可以用其他不同于样本均值的值来拓展$\{X_t\}$，例如，如果用线性外推的方法来拓展时间序列，就仍然可以用式(143)得到多分辨分析，但是用上面的方法得到方差分析的过程就不再适用了.）

第二种处理一般的长为 N 的序列的方法是将序列截断，得到一个长为2^{J_0}的整数倍的序列，这样我们就能用离散小波变换来处理了．如果令N''是比 N 小的且是2^{J_0}的整数倍的最大的整数，则可以定义两个缩短的序列

$$\mathbf{X}^{(1)} \equiv [X_0,\cdots,X_{N''-1}]^{\mathrm{T}},\quad \mathbf{X}^{(2)} \equiv [X_{N-N''},\cdots,X_{N-1}]^{\mathrm{T}},$$

两个序列的长度都为N''．令$\mathbf{W}_j^{(l)}$和$\mathbf{V}_{J_0}^{(l)}$表示$\mathbf{X}^{(l)}$的离散小波变换系数的各子向量，我们可以将两个分析结合起来按如下方法得到一个不精确的方差分析——但是却是一个精确的多分辨分析．由于有

$$\sum_{t=0}^{N''-1} X_t^2 = \sum_{j=1}^{J_0} \| \mathbf{W}_j^{(1)} \|^2 + \| \mathbf{V}_{J_0}^{(1)} \|^2$$

和

$$\sum_{t=N-N''}^{N-1} X_t^2 = \sum_{j=1}^{J_0} \| \mathbf{W}_j^{(2)} \|^2 + \| \mathbf{V}_{J_0}^{(2)} \|^2,$$

我们可以将上面两式进行平均得到

$$\begin{aligned}&\frac{1}{2}\sum_{t=0}^{N-N''-1} X_t^2 + \sum_{t=N-N''}^{N''-1} X_t^2 + \frac{1}{2}\sum_{t=N''}^{N-1} X_t^2\\ &= \sum_{j=1}^{J_0} \frac{1}{2}\left(\| \mathbf{W}_j^{(1)} \|^2 + \| \mathbf{W}_j^{(2)} \|^2 \right) + \frac{1}{2}\left(\| \mathbf{V}_{J_0}^{(1)} \|^2 + \| \mathbf{V}_{J_0}^{(2)} \|^2 \right).\end{aligned}$$

这导致了基于尺度的加权平方和的分解，其中半权重加在序列前面和后面的 $N-N''$个上. 为了得到相应的多分辨分析，我们形成

$$\mathbf{X}^{(l)} = \sum_{j=1}^{J_0} \mathcal{D}_j^{(l)} + \mathcal{S}_{J_0}^{(l)}, \quad l = 1,2,$$

然后结合对 X_0，…，$X_{N-N''-1}$ 的逐点分解(由 $\mathbf{X}^{(1)}$ 的多分辨分析给出)和 $X_{N''}$，…，X_{N-1} 的逐点分解(由 $\mathbf{X}^{(2)}$ 的多分辨分析给出)，以及再对 $X_{N-N''}$，…，$X_{N''-1}$ 平均两个逐点分解. 数学上，我们可以将得到的多分辨分析写为

$$\mathbf{X} = \sum_{l=1}^{2} \left(\sum_{j=1}^{J_0} I_{N,N''}^{(l)} \mathcal{D}_j^{(l)} + I_{N,N''}^{(l)} \mathcal{S}_{J_0}^{(l)} \right),$$

其中 $I_{N,N''}^{(l)}$ 是 $N \times N''$阶矩阵，其元素通过下面的例子来说明，例如 $N=6$，$N''=4$ 时，

$$I_{6,4}^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad I_{6,4}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(一个对上式稍稍修改的方案是结合$[X_l, \cdots, X_{N''-1+l}]^T$ 的 $N-N''+1$ 个分解，$l=0$，…，$N-N''$，这就与 Coifman 和 Donoho 在 1995 年提出的"循环旋转"(cycle spinning)非常接近了.)

第三个方法是修改塔式算法的公式，每一层的尺度系数多保留一个系数. 为了看清楚具体的操作方式，假定我们有 $V_{j-1,t}$，$t=0$，…，$M-1$，其中 M 为奇数. 我们应用通常的塔式算法来调整 $V_{j-1,t}$，$t=0$，…，$M-2$，以获得 $V_{j,t}$ 和 $W_{j,t}$，$t=0$，…，$\frac{M-1}{2}-1$. 于是分解的平方和形式为

$$\sum_{t=0}^{M-1} V_{j-1,t}^2 = \sum_{t=0}^{\frac{M-1}{2}-1} V_{j,t}^2 + \sum_{t=0}^{\frac{M-1}{2}-1} W_{j,t}^2 + V_{j-1,M-1}^2;$$

即最后的 $V_{j-1,M-1}^2$ 在左边单独留下了，不参与下一层的分解. 这个方案将 $\sum X_t^2$ 分解为小波系数和尺度系数的平方和(然而，尺度系数现在不再是在单个尺度上了：对于 J_0 层分解，从 $j=1$ 到 J_0-1 至多可以有一个尺度系数). 逆算法从 $\{W_{j,t}\}$、$\{V_{j,t}\}$ 和各层保留的系数 $V_{j-1,M-1}$ 重构 $\{V_{j-1,t}\}$. 我们用类似的方式定义细节 $\mathcal{D}_j$：对 $\{W_{j,t}\}$ 应用逆变换，其他的 $\{V_{j,t}\}$，$\{W_{j-1,t}\}$，$\{W_{j-2,t}\}$，$\{W_{1,t}\}$ 都置

为零(每一个保留的尺度系数也都置为零). 以类似的方式，我们定义光滑$\mathcal{S}_{J_0}$为将低层的小波系数都置为零后(在这种情况下所有的低层保留的尺度系数都保留)再做逆变换的结果.

作为例子，我们考虑对长为37的时间序列X_0，…，X_{36}的$J_0=3$的部分离散小波变换. 第一层变换将保留X_{36}，并将X_0，…，X_{35}分解成$W_{1,t}$和$V_{1,t}$，$t=0$，…，17. 由于$\{V_{1,t}\}$的长度是偶数，没有必要保留第一层的任何尺度系数，将$\{V_{1,t}\}$分解成$W_{2,t}$和$V_{2,t}$，$t=0$，…，8. 由于$\{V_{2,t}\}$有9个值，我们保留$V_{2,8}$并分解$V_{2,0}$，…，$V_{2,7}$，得到$W_{3,t}$和$V_{3,t}$，$t=0$，1，2，3. 分解的平方和的形式为

$$\sum_{t=0}^{36} X_t^2 = \sum_{t=0}^{17} W_{1,t}^2 + \sum_{t=0}^{8} W_{2,t}^2 + \sum_{t=0}^{3} W_{3,t}^2 + \sum_{t=0}^{3} V_{3,t}^2 + V_{2,8}^2 + X_{36}^2.$$

若要得到$\mathcal{D}_1$，我们将$W_{1,t}$，$t=0$，…，17进行逆变换，$V_{1,t}$，$t=0$，…，17和X_{36}都被置为零；类似地，若要得到$\mathcal{S}_1$，我们对$V_{1,t}$，$t=0$，…，17，和X_{36}做逆变换，而$W_{1,t}$，$t=0$，…，17都被置为零. 这样得到的多分辨分析满足$\mathbf{X}=\mathcal{S}_1+\mathcal{D}_1$.(更高层的多分辨分析可以类似得到.)

- **部分离散小波变换的分解层数J_0的选择**

正如小波滤波器的选择，部分离散小波变换的分解层数J_0选取也取决于特定的应用. 在心电图的例子中，注意令$J_0=6$导致了和物理尺度$\lambda_6\Delta t=0.36$s相关联的尺度系数$\mathbf{V}_6$. 这个选择有效地将心电图分析所不关心的大尺度的波动("基线偏移")分离到了到尺度系数，使得心跳节奏在各尺度的体现由分解得到的小波系数给出. 在后面我们将要给出的例子中(特别地，参见5.8、5.9和5.10节)，我们会考虑实际的物理意义来选择部分离散小波变换的分解层数J_0.

我们也应该注意到小波滤波器的宽度L会对选择J_0产生影响. 由于大的L会导致长度L_j更大的高层小波滤波器和尺度滤波器. 通常的方案是选择J_0使得$L_{J_0}\leqslant N<L_{J_0+1}$. 这个选择保证至少某些$\mathbf{W}_{J_0}$和$\mathbf{V}_{J_0}$系数不会被边界条件影响.

4.11节的评论与扩展

[1]这里，我们定量地研究循环假设对离散小波变换系数的影响. 注意第j层系数可以表示为

$$W_{j,t} = 2^{j/2}\widetilde{W}_{j,2^j(t+1)-1},\quad t=0,\cdots,N_j-1\equiv\frac{N}{2^j}-1,$$

其中，对于$t=0$，…，$N-1$，

$$2^{j/2}\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} h_{j,l}X_{t-l \bmod N}$$

(参见式(96c)和(96d)). 从练习[96]我们知道，$\{h_{j,l}\}$具有宽度$L_j\equiv(2^j-1)(L-1)+1$. 因此，在用$\{h_{j,l}\}$对$\{X_t\}$进行循环滤波得到了$\{2^{j/2}\widetilde{W}_{j,t}\}$，输出值的指标为$t=$

0，…，L_j-2 是用了循环假设，而指标为 $t=L_j-1$，…，$N-1$ 的值没有被循环假设影响. 第一个没有被循环假设影响的小波系数 $W_{j,t}$ 的指标为满足下式的最小的 t：

$$2^j(t+1)-1 \geqslant L_j-1,$$

或者等价地(经过一些约简)

$$t \geqslant (L-2)\left(1-\frac{1}{2^j}\right), \quad 即\ t=\left\lceil (L-2)\left(1-\frac{1}{2^j}\right)\right\rceil \equiv L_j', \tag{146a}$$

其中$\lceil x \rceil$表示大于或等于 x 的最小整数. 注意

$$L_1'=\frac{L}{2}-1, \quad 且 \quad \frac{L}{2}-1 \leqslant L_j' \leqslant L-2. \tag{146b}$$

进而，因为 $L_j'=L-2$ 成立当且仅当

$$(L-2)\left(1-\frac{1}{2^j}\right) > L-3,$$

我们看到 $L_j'=L-2$ 对所有的满足 $2^j>L-2$ 的 j 成立.

在假设 $L_j' \leqslant N_j=N/2^j$ 下，我们可以将 L_j'解释成直接被循环假设影响的第 j 层的小波系数 $W_{j,t}$，更一般地，受循环影响的系数的个数为 $\min\{L_j', N_j\}$. 我们称这些系数为边界系数. 注意 $\widetilde{W}_{j,t}(t=L_j-1, \cdots, N-1)$没有受循环影响. 我们对得到的序列下采样后得到小波系数 $\mathbf{W}_j$ 的元素，只有前 L_j'个系数为边界系数. 表 136 对滤波器宽 $L=2$，…，20 给出了 L_j'.

例如，对于 $j=1$ 和 $L=4$，只有式(81b)中的矩阵$\mathcal{W}_1$ 的第一行受到循环的影响，因此只有 $W_{1,0}$是边界系数. 表 136 精确给出了对于 $L=4$，$L_1'=1$.

一个类似的讨论说明 J_0 层部分离散小波变换边界尺度系数数目为 $\min\{L_{J_0}', N_{J_0}\}$. 然而练习[97]暗示，当 $N=2^J$，$J_0=J$ 时，对于完全离散小波变换情况不太一样，这时的完全离散小波变换的尺度系数和时间序列的样本均值成比例，不受循环的影响.

[2]当我们用最接近对称或者 coiflet 小波滤波器分解时，用得到的系数为纵坐标，实际时间为横坐标绘图(就像 4.8 节所描述的那样)，我们要决定 $\mathbf{W}_j$ 中的 L_j'个边界系数. 对于某个特定的最接近对称或 coiflet 小波滤波器，若我们将循环前移考虑进去，$W_{j,t}$从左到右的绘图顺序通常和它们在 $\mathbf{W}_j$ 中的顺序是不同的. 为了看清楚这一点，考虑 $j=1$ 的情形. 根据 4.8 节的讨论(为了方便假设 $t_0=0$，$\Delta t=1$)，我们得到 $W_{1,t}$相应于实际时间 $2t+1-|\nu_1^{(H)}| \bmod N$，$t=0$，…，$\frac{N}{2}-1$，对于最接近对称滤波器，我们可以用式(114c)写出

$$|\nu_1^{(H)}|=\frac{L}{2}+\delta, \quad 其中\ \delta=\begin{cases}0, & L/2\ 是偶数; \\ -1, & L=10\ 或\ 18; \\ 1, & L=14.\end{cases}$$

对于 coiflet 小波滤波器，我们有 $|\nu_1^{(H)}|=\frac{L}{3}$. 由于实际中 L 和 N 相比非常小，第一个最接近对称或者 coiflet 单位尺度上的系数 $W_{1,0}$ 对应的实际时间分别为

$$1-\frac{L}{2}-\delta \bmod N = N+1-\frac{L}{2}-\delta \quad \text{或者 } 1-\frac{L}{3} \bmod N = N+1-\frac{L}{3},$$

因此 $W_{1,0}$ 出现在实际时间段的最后部分. 系数 $W_{1,t}$ 相应的是实际时间的最前段，对于最接近对称小波滤波器，最小的指标 t 满足

$$2t+1-\frac{L}{2}-\delta \geqslant 0; \quad \text{即 } t=\left\lceil\frac{L+2\delta-2}{4}\right\rceil \equiv \gamma_1^{(H)},$$

或者，对于 L 总是 6 的倍数的 coiflet 滤波器来说，

$$2t+1-\frac{L}{3} \geqslant 0; \quad \text{即 } t=\left\lceil\frac{L-3}{6}\right\rceil=\frac{L}{6} \equiv \gamma_1^{(H)}.$$

$\mathbf{W}_1$ 的元素按绘图的顺序排列的序列实际上是

$$\mathcal{T}^{-\gamma_1^{(H)}}\mathbf{W}_1=[W_{1,\gamma_1^{(H)}}, W_{1,\gamma_1^{(H)}+1}, \cdots, W_{1,\frac{N}{2}-1}, W_{1,0}, \cdots, W_{1,\gamma_1^{(H)}-1}]^{\mathrm{T}}.$$

现在 $\mathbf{W}_1$ 中的前 $L_1'=\frac{L}{2}-1$ 个系数是边界系数，即指标为 $t=0, \cdots, \frac{L}{2}-2$ 的系数. 由于 $\mathcal{T}^{-\gamma_1^{(H)}}$ 的第一个元素的指标为 $\gamma_1^{(H)}$，且对所有的最接近对称和 coiflet 滤波器来说 $\gamma_1^{(H)} \leqslant \frac{L}{2}-2<L_1'$. 由此得出 $\mathcal{T}^{-\gamma_1^{(H)}}\mathbf{W}_1$ 的最后 $\gamma_1^{(H)}$ 个系数都被循环所影响，同样前 $\frac{L}{2}-1-\gamma_1^{(H)} \equiv \tilde{\gamma}_1^{(H)}$ 个系数也被循环所影响. 对更高尺度的类似分析表明，在假设 $L_j' \leqslant N_j$ 下，$\mathbf{W}_j$ 中元素绘图的顺序为 $\mathcal{T}^{-\gamma_j^{(H)}}\mathbf{W}_j$，对于最接近对称和 coiflet 小波滤波器我们分别有

$$\gamma_j^{(H)} \equiv \left\lceil\frac{L_j+2\delta+2}{2^{j+1}}-1\right\rceil \quad \text{和} \quad \gamma_j^{(H)} \equiv \left\lceil\frac{3L_j-L+6}{3 \cdot 2^{j+1}}-1\right\rceil.$$

对于这两类滤波器，我们总有 $0<\gamma_j^{(H)}<L_j'$. 因此 $\mathbf{W}_j$ 的前 $\gamma_j^{(H)}$ 个系数相应于实际时间的最后 $\gamma_j^{(H)}$ 个时间点，且都是边界系数. 另外有 $\tilde{\gamma}_j^{(H)} \equiv L_j'-\gamma_j^{(H)}$ 个边界系数，它们相应于实际时间的前 $\tilde{\gamma}_j^{(H)}$ 个时间点. 最接近对称的 $\tilde{\gamma}_j^{(H)}$ 和 $\gamma_j^{(H)}$ 的值在表 137a 中给出. 对尺度系数进行类似分析我们可以得到类似的量 $\gamma_{J_0}^{(G)}$ 和 $\tilde{\gamma}_{J_0}^{(G)}$，表 137b 给出了它们的值. 练习[4.27]就是要对 coiflet 滤波器计算类似的表.

[3]这里，我们考虑循环对细节的影响. 可以写成

$$\mathcal{D}_j=\mathcal{W}_j^{\mathrm{T}}\mathbf{W}_j=\sum_{n=0}^{N_j-1} W_{j,n}\mathcal{W}_{j,n\bullet},$$

其中$\mathcal{W}_{j,n\cdot}$是一个N维列向量，其元素构成了$\mathcal{W}_j$的第n行．当$L_j'\leqslant N_j$时，我们知道边界系数是$W_{j,n}$，$n=0$，…，$L_j'-1$，这些系数继续通过相应的向量$\mathcal{W}_{j,n\cdot}$的非零部分影响$\mathcal{D}_j$，即可以写成

$$\mathcal{D}_j=\sum_{n=0}^{L_j'-1}W_{j,n}\mathcal{W}_{j,n\cdot}+\sum_{n=L_j'}^{N_j-1}W_{j,n}\mathcal{W}_{j,n\cdot},$$

其中第一项和显然依赖于循环，但是第二项不依赖于循环．例如，考虑$j=1$和$L=4<N$的情形，这时只有$W_{1,0}$是边界系数．我们从式(81b)推出$\mathcal{D}_1$通过下式受循环影响：

$$\sum_{n=0}^{L_1'-1}W_{1,n}\mathcal{W}_{1,n\cdot}=W_{1,0}\mathcal{W}_{1,0\cdot}=W_{1,0}[h_1,h_0,0,\cdots,0,h_3,h_2]^{\mathrm{T}}.$$

因此$\mathcal{D}_1$的边界元素的指标为$n=0$，1，$N-2$和$N-1$．对于$j=1$和$4\leqslant L<N$，得到$L_1'=\frac{L}{2}-1>0$，边界系数$W_{1,0}$，$W_{1,1}$，…，$W_{1,L_1'-1}$通过下式影响$\mathcal{D}_1$：

$$\sum_{n=0}^{L_1'-1}W_{1,n}\mathcal{W}_{1,n\cdot}=W_{1,0}\begin{bmatrix}h_1\\h_0\\0\\\vdots\\0\\h_{L-1}\\\vdots\\h_2\end{bmatrix}+\cdots+W_{1,L_1'-1}\begin{bmatrix}h_{L-3}\\\vdots\\h_0\\0\\\vdots\\0\\h_{L-1}\\h_{L-2}\end{bmatrix}.$$

注意受到影响的元素的指标可以从上面的和式中的第一项和最后一项的非零部分得到：因为$\mathcal{W}_{1,n\cdot}$是通过对$\mathcal{W}_{1,0\cdot}$循环平移$2n$位得到的，对于$\mathcal{W}_{1,1\cdot}$，…，$\mathcal{W}_{1,L_1'-2\cdot}$的每个非零元素，在$\mathcal{W}_{1,0\cdot}$或者$\mathcal{W}_{1,L_1'-1\cdot}$中对应一个非零元素．在$\mathcal{D}_1$中受到影响的元素的指标为$n=0$，…，$L-3$(这可由$\mathcal{W}_{1,L_1'-1\cdot}$推出)以及$n=N-(L-2)$，…，$N-1$(从$\mathcal{W}_{1,0\cdot}$推出)．

以类似的方式，当$0<L_j'\leqslant N_j$时，可以推出$\mathcal{D}_j$中受到影响的元素的指标能通过$\mathcal{W}_{j,0\cdot}$和$\mathcal{W}_{j,L_j'-1\cdot}$的非零元素的指标决定．回忆$\langle\mathbf{X},\mathcal{W}_{j,0\cdot}\rangle=W_{j,n}$，我们可以用式(96d)得到

$$\langle\mathbf{X},\mathcal{W}_{j,0\cdot}\rangle=\sum_{l=0}^{L_j-1}h_{j,l}X_{2^j-1-l\bmod N}$$

及

$$\langle\mathbf{X},\mathcal{W}_{j,L_j'-1\cdot}\rangle=\sum_{l=0}^{L_j-1}h_{j,l}X_{2^jL_j'-1-l\bmod N}.$$

我们要求的指标就是上面两个和式中的 X_t 的指标. 从第一个式子得到指标

$$t=0,\cdots,2^j-1 \quad \text{和} \quad t=N+2^j-L_j,\cdots,N-1; \tag{149a}$$

从第二个式子我们得到

$$t=0,\cdots,2^jL'_j-1 \quad \text{和} \quad t=N+2^jL'_j-L_j,\cdots,N-1. \tag{149b}$$

由于 $L'_j\geqslant 1$，这两个指标集合的并推出了式(139)的结果(类似的结果对 $\mathcal{S}_{J_0}$ 也成立，$J_0=J$ 是特殊情形).

练习[149] 如果在 $\mathcal{D}_j$ 的最后部分有受到循环影响的值，我们必有式(149a)中的 L_j 满足 $L_j>2^j$，或者式(149b)中有 $L_j>2^jL'_j$. 证明，事实上在我们的假设 $L'_j>0$ 下这两个不等式一定成立. ◁

[4]我们应该注意在各种文献中基于离散 Daubechies 小波的离散小波变换的变换$\mathcal{W}$的定义是不唯一的，因此，如果本书中的某幅图要和一个小波软件包的计算结果相比较，我们要注意这个软件包中小波的实现是基于什么定义的. 例如，Bruce and Gao(1996a)定义的哈尔离散小波变换和我们的定义是不同的，他们的小波变换的矩阵$\mathcal{W}$与本书定义的不同，他们定义的小波变换阵为

$$\begin{bmatrix}\mathcal{V}_J\\ \mathcal{W}_J\\ \vdots\\ \mathcal{W}_2\\ \mathcal{W}_1\end{bmatrix},\quad \text{而不是我们定义的}\mathcal{W}=\begin{bmatrix}\mathcal{W}_1\\ \mathcal{W}_2\\ \vdots\\ \mathcal{W}_J\\ \mathcal{V}_J\end{bmatrix}.$$

这两个定义本质上是一样的，只是变换以后的哈尔小波系数差一个反序.

对于除哈尔之外的其他小波，这两种差异就没这么简单了. 例如，图 150 给出了对于 Bruce and Gao(1996a)给出的 $N=16$ 的 D(4)离散小波变换矩阵，图 60 是我们定义的离散小波变换矩阵，仔细比较这两幅图我们可以看出，即使我们不考虑顺序，这两个变换矩阵的每一行都差一个平移(从图 60 到图 150 我们需要将图 60 的每一行向右平移 $2\tau_j-1$ 位). 因此，这两个 D(4)小波变换的离散小波变换系数是不一致的，除了最后一个和样本均值成正比的系数!

导致小波变换这种差异的原因可以找出来. 例如，小波变换内在的有下采样(或上采样)过程，执行下采样的时候我们可以选择保留偶数指标的系数或者奇数指标的系数(事实上时间序列有时用 1，…，N 来作为指标而不是我们这里选择的 0，…，$N-1$). 此外，我们用滤波的过程来执行小波变换(即卷积)，但是这也等价于做点积(即自相关). 这种选择很大程度上和各领域的习俗有关：电子工程学者倾向于将小波变换作为一个滤波过程来执行，而数学家倾向于用点积的观点来看小波变换. 还没有解决这种差异的根本的方法，因此初学者要小心这一点!

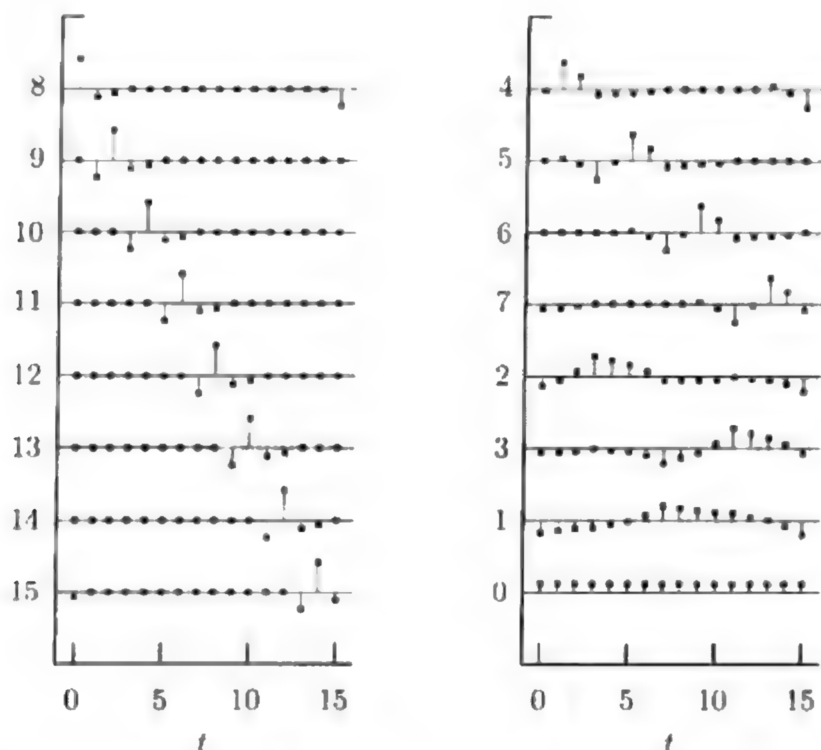


图 150 基于 Bruce and Gao(1996a)定义的 D(4) 离散小波变换的变换矩阵的行向量，参见图 60

4.12 小结

令 $\mathbf{X}$ 是一个 N 维向量，其元素是实值的时间序列 $\{X_t: t=0, \cdots, N-1\}$，其中样本的长度 N 为 2^{J_0} 的整数倍. $\mathbf{X}$ 的 J_0 层部分离散小波变换是一个规范正交变换 $\mathbf{W}=\mathcal{W}\mathbf{X}$，其中 $\mathbf{W}$ 是元素为离散小波变换系数的 N 维向量，$\mathcal{W}$ 是 $N\times N$ 阶实值矩阵，由这个矩阵定义了离散小波变换(如果 $N=2^J$ 且 $J_0=J$，我们就得到完全离散小波变换). 离散小波变换系数 $\mathbf{W}$ 和小波变换矩阵 $\mathcal{W}$ 可以分块成

$$\mathbf{W}=\begin{bmatrix}\mathbf{W}_1\\ \mathbf{W}_2\\ \vdots\\ \mathbf{W}_{J_0}\\ \mathbf{V}_{J_0}\end{bmatrix} \quad \text{且} \quad \mathcal{W}=\begin{bmatrix}\mathcal{W}_1\\ \mathcal{W}_2\\ \vdots\\ \mathcal{W}_{J_0}\\ \mathcal{V}_{J_0}\end{bmatrix}, \tag{150}$$

因此 $\mathbf{W}_j=\mathcal{W}_j\mathbf{X}$ 和 $\mathbf{V}_{J_0}=\mathcal{V}_{J_0}\mathbf{X}$. 这里 $\mathbf{W}_j$ 是尺度 $\tau_j\equiv2^{j-1}$ 上的 $N_j\equiv N/2^j$ 维的小波系数构成的向量，$\mathcal{W}_j$ 是一个 $N_j\times N$ 维矩阵；$\mathbf{V}_{J_0}$ 是在尺度 $\lambda_{J_0}\equiv2^{J_0}$ 上的尺度系数构成的 N_{J_0} 维向量；$\mathcal{V}_{J_0}$ 是 $N_{J_0}\times N$ 维矩阵. 向量 $\mathbf{X}$ 可以通过下式由 $\mathbf{W}$ 重构：

$$\mathbf{X}=\mathcal{W}^{\mathrm{T}}\mathbf{W}=\sum_{j=1}^{J_0}\mathcal{W}_j^{\mathrm{T}}\mathbf{W}_j+\mathcal{V}_{J_0}^{\mathrm{T}}\mathbf{V}_{J_0}\equiv\sum_{j=1}^{J_0}\mathcal{D}_j+\mathcal{S}_{J_0},$$

上式定义了 $\boldsymbol{X}$ 的一个多分辨分析(MRA)，即将 $\boldsymbol{X}$ 分解为 N 维向量 $\mathcal{D}_j \equiv \mathcal{W}_j^{\mathrm{T}} \boldsymbol{W}_j$(第 j 层细节)和 $\mathcal{S}_{J_0} \equiv \mathcal{V}_{J_0}^{\mathrm{T}} \boldsymbol{V}_{J_0}$(第 J_0 层光滑)的和，每个分量都和一个特定的尺度($\mathcal{D}_j$ 对应尺度 τ_j，$\mathcal{S}_{J_0}$ 对应尺度 λ_{J_0})相关. 一个尺度-尺度的方差分析能基于能量的分解

$$\|\boldsymbol{X}\|^2 = \|\boldsymbol{W}\|^2 = \sum_{j=1}^{J_0} \|\boldsymbol{W}_j\|^2 + \|\boldsymbol{V}_{J_0}\|^2 = \sum_{j=1}^{J_0} \|\mathcal{D}_j\|^2 + \|\mathcal{S}_{J_0}\|^2.$$

在实际应用中，离散小波变换矩阵 $\mathcal{W}$ 不是直接求得的，我们通过塔式算法来计算 $\boldsymbol{W}$，计算过程需要用到小波滤波器和尺度滤波器(这个算法实际上明确地定义了离散小波变换). 由定义，宽为偶数 L 的滤波器 $\{h_l: l=0, \cdots, L-1\}$(隐含 $h_0 \neq 0$，$h_{L-1} \neq 0$)称为小波滤波器，如果

$$\sum_{l=0}^{L-1} h_l = 0 \quad 且 \quad \sum_{l=0}^{L-1} h_l h_{l+2n} = \begin{cases} 1, & n = 0; \\ 0, & n \text{ 是非零整数}. \end{cases}$$

其中第二个和式表示小波滤波器的规范正交性(在上式中，$h_l \equiv 0$，$l<0$ 或 $l \geqslant L$，因此我们实际上考虑 $\{h_l\}$ 是一个至多有 L 个非零元素的无限长的序列). 尺度滤波器借助于小波滤波器来定义，这两个滤波器之间需要满足"正交镜像"的关系

$$g_l \equiv (-1)^{l+1} h_{L-1-l}.$$

这个滤波器满足条件：

$$\sum_{l=0}^{L-1} g_l g_{l+2n} = \begin{cases} 1, & n = 0 \\ 0, & \text{其他}; \end{cases} \quad 且 \quad \sum_{l=0}^{L-1} g_l h_{l+2n} = 0 \text{ 对所有的 } n \text{ 成立}.$$

不失一般性，我们还可以假设 $\sum_l g_l = \sqrt{2}$. 如果令 $H(\cdot)$ 是 $\{h_l\}$ 的传递函数，即

$$H(f) \equiv \sum_{l=-\infty}^{\infty} h_l \mathrm{e}^{-\mathrm{i}2\pi fl} = \sum_{l=0}^{L-1} h_l \mathrm{e}^{-\mathrm{i}2\pi fl},$$

并且如果我们用 $\mathcal{H}(f) \equiv |H(f)|^2$ 定义相应的平方增益函数，那么小波滤波器的正交性等价于

$$\mathcal{H}(f) + \mathcal{H}\left(f + \frac{1}{2}\right) = 2 \quad \text{对所有的 } f \text{ 成立}.$$

如果我们令 $G(\cdot)$ 和 $\mathcal{G}(\cdot)$ 是尺度滤波器的传递函数和平方增益函数，则有 $\mathcal{G}(f) = \mathcal{H}\left(\frac{1}{2} - f\right)$，从中可以得到

$$\mathcal{G}(f) + \mathcal{G}\left(f + \frac{1}{2}\right) = 2 \quad 且 \quad \mathcal{G}(f) + \mathcal{H}(f) = 2 \text{ 对所有的 } f \text{ 成立}.$$

在实际应用中，$\{h_l\}$ 是一个高通滤波器，其通带为 $\frac{1}{4} \leqslant |f| \leqslant \frac{1}{2}$，而 $\{g_l\}$ 是一

个低通滤波器，其标称通带为 $0 \leqslant |f| \leqslant \frac{1}{4}$. 由于这两个滤波器的通带都是满带的一半，因此 $\{h_l\}$ 和 $\{g_l\}$ 可以称为是“半带”滤波器.

定义了 $\{h_l\}$ 和 $\{g_l\}$ 之后，第 j 层塔式算法就是对 N_{j-1} 个元素进行循环滤波，这 N_{j-1} 个元素为

$$\mathbf{V}_{j-1} \equiv [V_{j-1,0}, V_{j-1,1}, \cdots, V_{j-1,N_{j-1}-1}]^{\mathrm{T}},$$

然后保留滤波输出值的奇数指标的值，这产生了第 j 层小波系数和尺度系数，即

$$W_{j,t} \equiv \sum_{l=0}^{L-1} h_l V_{j-1,2t+1-l \bmod N_{j-1}}, \quad V_{j,t} \equiv \sum_{l=0}^{L-1} g_l V_{j-1,2t+1-l \bmod N_{j-1}},$$

$t=0$，…，N_j-1（这些系数是 $\mathbf{W}_j$ 和 $\mathbf{V}_j$ 的元素）. 每隔一位保留一个滤波值的过程称为因子为 2 的下采样. 令 $\mathbf{V}_0 \equiv \mathbf{X}$，我们从 $j=1$ 开始塔式算法，经过 $j=2$，3，…，J_0，我们得到的所有向量可以形成小波系数矩阵 $\mathcal{W}$，即 J_0 个小波系数向量 $\mathbf{W}_1$，…，$\mathbf{W}_{J_0}$ 和单个尺度系数向量 $\mathbf{V}_{J_0}$（其余的尺度系数向量 $\mathbf{V}_1$，…，$\mathbf{V}_{J_0-1}$ 可以认为是计算的中间量）.

虽然在实际中 $\mathbf{W}_j$ 和 $\mathbf{V}_j$ 是用塔式算法计算出来的，但是理论上我们可以直接利用下式从 $\mathbf{X}$ 得到 $\mathbf{W}_j$ 和 $\mathbf{V}_j$：

$$W_{j,t} = \sum_{l=0}^{L_j-1} h_{j,l} X_{2^j(t+1)-1-l \bmod N}, \quad V_{j,t} = \sum_{l=0}^{L_j-1} g_{j,l} X_{2^j(t+1)-1-l \bmod N},$$

其中 $\{h_{j,l}\}$ 和 $\{g_{j,l}\}$ 是第 j 层等价小波滤波器和尺度滤波器，它们的长度都是 $L_j \equiv (2^j-1)(L-1)+1$（这里 $h_{1,l} \equiv h_l$ 和 $g_{1,l} \equiv g_l$）. 它们的传递函数为

$$H_j(f) \equiv H(2^{j-1}f) \prod_{l=0}^{j-2} G(2^l f), \quad G_j(f) \equiv \prod_{l=0}^{j-1} G(2^l f)$$

（再次 $H_1(f) \equiv H(f)$ 和 $G_1(f) \equiv G(f)$）. 滤波器 $\{h_{j,l}\}$ 是个带通滤波器，其标称通带为 $1/2^{j+1} \leqslant |f| \leqslant 1/2^j$，而 $\{g_{j,l}\}$ 是个低通滤波器，标称通带为 $0 \leqslant |f| \leqslant 1/2^{j+1}$. 表 154 总结了小波滤波器和尺度滤波器以及由它们产生的许多重要性质.

给定 $\mathbf{W}_j$ 和 $\mathbf{V}_j$，我们可以通过第 j 步逆塔式算法重构(合成) $\mathbf{V}_{j-1}$ 的元素，即

$$V_{j-1,t} = \sum_{l=0}^{L-1} h_l W^{\uparrow}_{j,t+l \bmod N_{j-1}} + \sum_{l=0}^{L-1} g_l V^{\uparrow}_{j,t+l \bmod N_{j-1}},$$

$t=0$，1，…，$N_{j-1}-1$，其中

$$W^{\uparrow}_{j,t} \equiv \begin{cases} 0, & t = 0,2,\cdots,N_{j-1}-2; \\ W_{j,\frac{t-1}{2}}, & t = 1,3,\cdots,N_{j-1}-1. \end{cases}$$

($V_{j,t}^{\uparrow}$也类似地定义). 在$\{W_{j,t}\}$的每两个元素中插入一个零形成$\{W_{j,t}^{\uparrow}\}$的过程称为是因子为 2 的上采样.

构成塔式算法的滤波运算完全决定了式(150)中定义离散小波变换的 $N\times N$ 阶矩阵$\mathcal{W}$. 回忆式(150)中对$\mathcal{W}$的分块，可以写成

$$\mathcal{W}_j = \mathcal{B}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1 \quad 和 \quad \mathcal{V}_{J_0} = \mathcal{A}_{J_0} \mathcal{A}_{J_0-1} \cdots \mathcal{A}_1,$$

其中$\mathcal{B}_j$和$\mathcal{A}_j$是$N_j\times N_{j-1}$阶矩阵，满足$\mathcal{B}_j \mathcal{A}_j^{\mathrm{T}} = \mathcal{A}_j \mathcal{B}_j^{\mathrm{T}} = 0_{N_j}$和$\mathcal{B}_j \mathcal{B}_j^{\mathrm{T}} = \mathcal{A}_j \mathcal{A}_j^{\mathrm{T}} = I_{N_j}$(这里$0_{N_j}$是$N_j\times N_j$阶矩阵，它所有的元素都为零). $\mathcal{B}_j$的第一行为:

$$[h_1^{\circ}, h_0^{\circ}, h_{N_{j-1}-1}^{\circ}, h_{N_{j-1}-2}^{\circ}, \cdots, h_3^{\circ}, h_2^{\circ}]^{\mathrm{T}},$$

其中$\{h_l^{\circ}\}$是以长 N_{j-1} 为周期的周期化后得到的序列$\left(因此满足\{h_l^{\circ}\}\leftrightarrow\left\{H\left(\frac{k}{N_{j-1}}\right)\right\}\right)$. 其后的行是通过将第一行向右进行跨度为 $2k(k=1,\cdots,N_j-1)$的循环平移; 特别地，最后一行是

$$[h_{N_{j-1}-1}^{\circ}, h_{N_{j-1}-2}^{\circ}, \cdots, h_1^{\circ}, h_0^{\circ}]^{\mathrm{T}}.$$

类似地可以构造$\mathcal{A}_j$，只要将上述过程中的$\{h_l\}$用$\{g_l\}$代替.

$\mathcal{W}_j$ 的各行可以直接通过第 j 层小波滤波器$\{h_{j,l}\}$以 N 为周期进行周期化得到. 周期化后的滤波器记为$\{h_{j,l}^{\circ}\}$，并且满足

$$\{h_{j,l}^{\circ}\}\leftrightarrow\left\{H\left(\frac{2^{j-1}k}{N}\right)G\left(\frac{2^{j-2}k}{N}\right)\cdots G\left(\frac{k}{N}\right)\right\}.$$

$\mathcal{W}_j$ 的第一行为

$$[h_{j,2^j-1}^{\circ}, h_{j,2^j-2}^{\circ}, \cdots, h_{j,1}^{\circ}, h_{j,0}^{\circ}, h_{j,N-1}^{\circ}, h_{j,N-2}^{\circ}, \cdots, h_{j,2^j+1}^{\circ}, h_{j,2^j}^{\circ}]^{\mathrm{T}}.$$

其后的各行可以通过对第一行向右循环平移 $k2^j(k=1,\cdots,N_j-1)$得到; 特别地，最后一行为

$$[h_{j,N-1}^{\circ}, h_{j,N-2}^{\circ}, \cdots, h_{j,1}^{\circ}, h_{j,0}^{\circ}]^{\mathrm{T}}.$$

类似地可以构造$\mathcal{V}_{J_0}$，只要将上述过程中的$\{h_{j,l}^{\circ}\}$用$\{g_{J_0,l}^{\circ}\}$代替即可.

小波滤波器定义中的性质(求和为零和规范正交性质)不一定能导出一个小波系数可以解释为某一特定尺度上的相邻加权平均的变化率的离散小波变换. 通过考虑一些正则条件(参见 11.9 节)，Daubechies(1992，6.1 节)定义了一类很有用的小波滤波器，由这些小波滤波器得到的小波系数可以解释成某一特定尺度上的相邻加权平均的变化率. 由定义，宽 L 为偶数的 Daubechies 小波滤波器的平方增益函数为

$$\mathcal{H}^{(\mathrm{D})}(f) \equiv \mathcal{D}^{\frac{L}{2}}(f)\,\mathcal{A}_L(f),$$

表 154

$\{h_l\} \leftrightarrow H(\cdot)$	$\{g_l\} \leftrightarrow G(\cdot)$
$h_l = (-1)^l g_{L-1-l}$	$g_l \equiv (-1)^{l+1} h_{L-1-l}$
$H(f) = -\mathrm{e}^{-i2\pi f(L-1)} G\left(\frac{1}{2} - f\right)$	$G(f) = \mathrm{e}^{-i2\pi f(L-1)} H\left(\frac{1}{2} - f\right)$
$\sum_l h_l = H(0) \equiv 0$	$\sum_l g_l = G(0) = \sqrt{2}$
$\sum_l h_l^2 \equiv 1$	$\sum_l g_l^2 = 1$
$\sum_l h_l h_{l+2n} \equiv 0$，$n \neq 0$	$\sum_l g_l g_{l+2n} = 0$，$n \neq 0$
$\sum_l g_l h_{l+2n} = 0$	
$\mathcal{H}(f) \equiv \mid H(f) \mid^2$	$\mathcal{G}(f) \equiv \mid G(f) \mid^2$
$\mathcal{H}(f) + \mathcal{H}\left(f + \frac{1}{2}\right) = 2$	$\mathcal{G}(f) + \mathcal{G}\left(f + \frac{1}{2}\right) = 2$
$\mathcal{G}(f) + \mathcal{H}(f) = 2$	
$W_{1,t} \equiv \sum_l h_l X_{2t+1-l \bmod N}$	$V_{1,t} \equiv \sum_l g_l X_{2t+1-l \bmod N}$
$W_{j,t} \equiv \sum_l h_l V_{j-1,2t+1-l \bmod N_{j-1}}$	$V_{j,t} \equiv \sum_l g_l V_{j-1,2t+1-l \bmod N_{j-1}}$
$h_{1,l} \equiv h_l$，$H_1(f) \equiv H(f)$	$g_{1,l} \equiv g_l$，$G_1(f) \equiv G(f)$
$H_j(f) \equiv H(2^{j-1}f) \prod_{l=0}^{j-2} G(2^l f)$	$G_j(f) \equiv \prod_{l=0}^{j-1} G(2^l f)$
$H_j(f) = H(2^{j-1}f) G_{j-1}(f)$	$G_j(f) = G(2^{j-1}f) G_{j-1}(f)$
$\{h_{j,l}\} \leftrightarrow H_j(\cdot)$	$\{g_{j,l}\} \leftrightarrow G_j(\cdot)$
$\sum_l h_{j,l} = H_j(0) = 0$	$\sum_l g_{j,l} = G_j(0) = 2^{j/2}$
$\sum_l h_{j,l}^2 = 1$	$\sum_l g_{j,l}^2 = 1$
$\sum_l h_{j,l} h_{j,l+2^j n} = 0$，$n \neq 0$	$\sum_l g_{j,l} g_{j,l+2^j n} = 0$，$n \neq 0$
$\sum_l g_{j,l} h_{j,l+2^j n} = 0$	
$\mathcal{H}_j(f) \equiv \mid H_j(f) \mid^2$	$\mathcal{G}_j(f) \equiv \mid G_j(f) \mid^2$
$W_{j,t} = \sum_l h_{j,l} X_{2^j(t+1)-1-l \bmod N}$	$V_{j,t} = \sum_l g_{j,l} X_{2^j(t+1)-1-l \bmod N}$

注：小波滤波器和尺度滤波器的重要关系. 因为对于 $l<0$ 或者 $l \geq L$ 有 $h_l = g_l = 0$，因此对 h_l 或者 g_l 求和可以是从 $l=0$ 加到 $l=L-1$，也可以是从负无穷加到正无穷；类似地，对 $h_{j,l}$ 和 $g_{j,l}$ 的求和可以是从 $l=0$ 加到 $l=L_j-1$，也可以是从负无穷加到正无穷（注意 $L_j \equiv (2^j-1)(L-1)+1$ 且 $N_j \equiv N/2^j$）.

其中$\mathcal{D}(f)\equiv 4\sin^2(\pi f)$定义了差分滤波器$\{1, -1\}$的平方增益函数，并且

$$\mathcal{A}_L(f)\equiv\frac{1}{2^{L-1}}\sum_{l=0}^{\frac{L}{2}-1}\binom{\frac{L}{2}-1+l}{l}\cos^{2l}(\pi f)$$

构成了低通滤波器的平方增益函数. 可以证明(练习[106])$\mathcal{H}^{(D)}(f)+\mathcal{H}^{(D)}\left(f+\frac{1}{2}\right)=2$对所有的$f$成立，且$\mathcal{H}^{(D)}(0)=0$，这意味着$\sum h_l=0$，因此任何滤波器$\{h_l\}$若其平方增益函数为$\mathcal{H}^{(D)}(\cdot)$，则由定义这个滤波器一定是小波滤波器. 上面的对$\mathcal{H}^{(D)}(\cdot)$的分析表明 Daubechies 小波滤波器可以等价地看成是一个低通滤波器(加权平均运算)级联$\frac{L}{2}$个差分滤波器(导致了总体上的差分性质). 相应于 Daubechies 小波滤波器的尺度滤波器$\{g_l\}$具有平方增益函数

$$\mathcal{G}^{(D)}(f)\equiv\mathcal{H}^{(D)}\left(\frac{1}{2}-f\right)=2\cos^L(\pi f)\sum_{l=0}^{\frac{L}{2}-1}\binom{\frac{L}{2}-1+l}{l}\sin^{2l}(\pi f).$$

不失一般性，宽为$L=2$或者$L=4$的 Daubechies 小波滤波器可以分别取成哈尔小波滤波器$\left\{h_0=\frac{1}{\sqrt{2}},\ h_1=-\frac{1}{\sqrt{2}}\right\}$和式(59a)描述的 D(4)小波滤波器$\{h_0, h_1, h_2, h_3\}$. 一般说来，存在不止一个小波滤波器$\{h_l: l=0, \cdots, L-1\}$满足平方增益函数$\mathcal{H}^{(D)}(\cdot)$. 当$L$增加时，我们可以加一些额外的标准来选择唯一的小波滤波器，或者等价地说，来选择唯一的尺度滤波器. Daubechies(1992，6.4 节)讨论了这两个标准. 第一种标准是选择尺度滤波器$\{g_l^{(\mathrm{ep})}\}$的平方增益函数$\mathcal{G}^{(D)}(\cdot)$满足

$$\sum_{l=0}^{m}g_l^2\leqslant\sum_{l=0}^{m}[g_l^{(\mathrm{ep})}]^2\quad 对\ m=0,\cdots,L-1\ 都成立,$$

其中$\{g_l\}$是任意一个满足平方增益函数为$\mathcal{G}^{(D)}(\cdot)$的滤波器. 这个选择称为是“最小相位”选择，得到的滤波器具有最小延迟. 我们满足这个标准的滤波器是 D(L)滤波器，$L=2, 4, \cdots$. 第二个标准是选择尺度函数使得其传递函数$G(f)=[\mathcal{G}^{(D)}(f)]^{1/2}e^{i\theta^{(G)}(f)}$的相位函数$\theta^{(G)}(\cdot)$尽量地接近线性相位滤波器(这里通过式(112a)进行选择). 我们满足这个标准的滤波器为 LA(L)滤波器，$L=8, 10, \cdots$，其中“LA”表示“最接近对称”(least asymmetric). 最接近对称滤波器的好处是我们可以用$\tilde{\nu}$的值ν使得上式最小，来使尺度系数和小波系数对齐，这样它们就可以近似地看成是零相位滤波器的输出. 这个逼近零相位的性质非常重要，因为它允许我们将离散小波变换系数和原始的时间序列中发生的事件有意义地联系起来. 特别地，如果假设X_t相应的实际时间为$t_0+t\Delta t$，则逼近零相位的性质使得最接近对称离散小波变换的小波系数$W_{j,t}$相应的实际时间为

$$t_0+(2^j(t+1)-1-|\nu_j^{(H)}|\bmod N)\Delta t,\quad t=0,\cdots,N_j-1,$$

其中

$$|\nu_j^{(H)}|=\frac{L_j}{2}+\frac{L}{2}+\nu-1,$$

并且

$$\nu=\begin{cases}-\frac{L}{2}+1, & L=8,12,16 \text{ 或 } 20(\text{即 } L/2 \text{ 是偶数});\\ -\frac{L}{2}, & L=10 \text{ 或 } 18;\\ -\frac{L}{2}+2, & L=14.\end{cases}$$

对于尺度系数 $V_{j,t}$，通过将 $|\nu_j^{(H)}|$ 换成

$$|\nu_j^{(G)}|=\frac{L_j-1}{L-1}|\nu|,$$

我们得到了一个类似的表达式.

coiflet 小波滤波器——记为 C(L)，$L=6$，12，18，24 或 30——是不同于 Daubechies 小波滤波器的另一类滤波器，它比最接近对称滤波器更接近零相位；然而这些滤波器的曲线形状不太好，对于给定的宽度它具有数量更少的内在的差分运算($L/3$ 而不是 Daubechies 小波滤波器的 $L/2$). 我们可以将 coiflet 小波系数和尺度系数用一个类似于最接近对称滤波器情形时的公式和实际时间联系起来，其中我们用的 ν 为

$$\nu=-\frac{2L}{3}+1.$$

4.13 练习

[4.1] 使用式(59b)的条件，解出式(60)中的 a 和 b，并且证明其中一组解导出了我们在式(59a)给出的 h_0，h_1，h_2，h_3. 共有多少组解？它们相应的 h_0，h_1，h_2，h_3 是多少？

[4.2] 验证图 62 中的左上图给出的是对 $\{X_{1,t}\}$ 的哈尔小波变换(这个时间序列的值在图 42 的说明中给出了).

[4.3] 验证离散哈尔小波变换的式(63a)和(63b)给出经验功率谱.

[4.4] 见图 65 的左上图和左下图，分别对应图 42 序列 $\{X_{1,t}\}$ 和 $\{X_{2,t}\}$ 的细节 $\mathcal{D}_1$，是用 D(4)小波变换得到的. 我们已知这两个序列只在 $t=13$ 处不一样，解释为什么相应的细节 $\mathcal{D}_1$ 在 $t=10$，…，15 处不同. 如果用 $L=6$ 的小波滤波器，细节 $\mathcal{D}_1$ 在 t 为何值的时候不同？

[4.5] 对于哈尔小波滤波器 $\{h_0=1/\sqrt{2},\ h_1=-1/\sqrt{2}\}$ 验证式(69d).

[4.6] 假定 $\{h_l:\ l=0,\ \cdots,\ 11\}$ 是一个长为 $L=12$ 的小波滤波器，基于这个小波滤波器构造离散小波变换矩阵 $\mathcal{W}$. 如果 $N=4$，$\mathcal{W}_{0\cdot}^{T}$ 和 $\mathcal{W}_{1\cdot}^{T}$ 的前两个行向量用 h_l 表示是什么形式的？如果 $N=8$，这两行又是什么形式的？

[4.7] 当 $L>N$ 时 $\mathcal{W}_1$ 的各行是规范正交的，对于 $N=4$ 和 $L=6$ 构造一个具体的例子. 利用式(69b)和式(69c)直接验证 $\mathcal{W}_1\ \mathcal{W}_1^{T}=I_2$，其中 I_2 是 2×2 阶单位阵.

[4.8] 证明任何小波滤波器 $\{h_l\}$ 的平方增益函数都满足 $\mathcal{H}(0)=0$ 且和 $\mathcal{H}\left(\frac{1}{2}\right)=2$ (参见图 73 的第一列图).

[4.9] 令 $\boldsymbol{X}=[X_0,\ X_1,\ X_2,\ X_3]^{T}$，假定我们对 $\boldsymbol{X}$ 进行因子为 2 的下采样，得到 $\boldsymbol{X}_{\downarrow}$. 如果对 $\boldsymbol{X}_{\downarrow}$ 进行上 2 采样将得到什么样的向量？

[4.10] 基于式(90a)中定义的 $2\widetilde{W}_{2,t}$，解释为什么

$$2\widetilde{W}_{2,t} = \sum_{l=0}^{L_2-1} h_{2,l} X_{t-l \bmod N}, \quad t = 0, \cdots, N-1,$$

和为什么

$$W_{2,t} = \sum_{l=0}^{L_2-1} h_{2,l} X_{4(t+1)-1-l \bmod N}, \quad t = 0, \cdots, \frac{N}{4}-1.$$

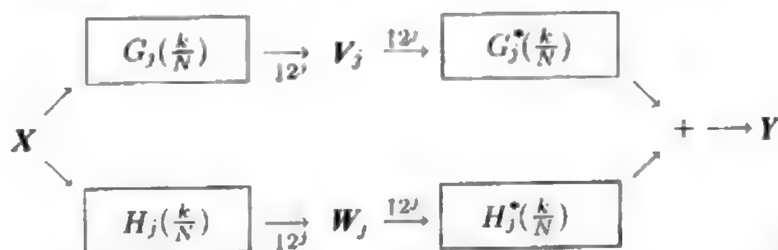
[4.11] 对于 D(4) 小波滤波器，计算用来对 $\{X_t\}$ 滤波分别得到 $\{W_{2,t}\}$ 和 $\{W_{3,t}\}$（尺度分别为 $\tau_2=2$ 和 $\tau_3=4$ 上的小波系数）的滤波器 $\{h_{2,l}\}$ 和 $\{h_{3,l}\}$ 的元素. 验证图 91 中绘的 $\{h_{2,l}\}$ 的平方增益函数的图形是正确的，并绘出 $\{h_{3,l}\}$ 的平方增益函数的图.

[4.12] 令 $\mathcal{H}_j(\cdot)$ 和 $\mathcal{G}_j(\cdot)$ 是分别相应于式(96b)和式(97)的平方增益函数. 求出 $\int_{-\frac{1}{2}}^{\frac{1}{2}} \mathcal{H}_j(f)\mathrm{d}f$ 和 $\int_{-\frac{1}{2}}^{\frac{1}{2}} \mathcal{G}_j(f)\mathrm{d}f$ 各等于什么?

[4.13] 验证用哈尔小波滤波器和尺度滤波器的塔式算法得到的变换与 4.1 节对 $N=16$ 给出的哈尔离散小波变换矩阵 $\mathcal{W}$ 是相同的. 特别地，对 $j=1$，2，3，4 计算 $\mathcal{B}_j$ 和 $\mathcal{A}_j$，并证明 $\mathcal{B}_2\mathcal{A}_1$，$\mathcal{B}_3\mathcal{A}_2\mathcal{A}_1$，$\mathcal{B}_4\mathcal{A}_3\mathcal{A}_2\mathcal{A}_1$ 和 $\mathcal{A}_4\mathcal{A}_3\mathcal{A}_2\mathcal{A}_1$ 与图 57 的右列的图给出的对应 $\mathcal{W}$ 的后 8 行是相同的.

[4.14] 绘出图 98b 中对应的 LA(8) 的第 j 层尺度滤波器和小波滤波器的平方增益函数 $\mathcal{G}_j^{(\mathrm{D})}(\cdot)$ 和 $\mathcal{H}_j^{(\mathrm{D})}(\cdot)$ 的图形.

[4.15] 令 $H_j(\cdot)$ 和 $G_j(\cdot)$ 是分别相应于第 j 层小波滤波器 $\{h_{j,l}\}$ 和尺度滤波器 $\{g_{j,l}\}$ 的传递函数. 在什么情况下下图中的 $\boldsymbol{Y}=\boldsymbol{X}$ 成立?



[4.16] 证明 $\mathcal{W}_1^{\mathrm{T}}\mathcal{W}_1+\mathcal{V}_1^{\mathrm{T}}\mathcal{V}_1=I_N$，此式可称为是 N 阶单位阵的 $j=1$ 阶分解. 给出并证明 $j>1$ 时的单位阵的分解形式.

[4.17] 根据 4.6 节评论与扩展中的[1]给出的伪代码，用你最熟悉的计算机编程语言写出离散小波变换塔式算法及其逆算法. 用 D(4) 小波滤波器和尺度滤波器（见式(59a)和(75d)）通过计算在图 42 说明中给出的 16 点时间序列 $\{X_{1,t}\}$ 的离散小波变换和逆变换来验证你写的程序的正确性.

[4.18] 对图 42 的说明中给出的 16 点时间序列 $\{X_{1,t}\}$ 用 D(4) 小波的塔式算法及其逆算法计算细节 $\mathcal{D}_3$.

[4.19] 对于 $t=0$，1，…，1 023，定义“线性鸣叫”时间序列

$$X_t = \cos\left(\frac{3\pi}{64}\left[1+\frac{t}{1\,023}\right]t+0.82\right) = \cos(2\pi f_t t + 0.82),$$

其中 $f_t \equiv \frac{3}{128}+\frac{3}{128}\left(\frac{t}{1\,023}\right)$，因此 f_t 的范围是从 $f_0=\frac{3}{128}$ 到 $f_{1\,023}=\frac{3}{64}$. 如果我们考虑到上式中余弦函数的参数是 t 的连续函数，就可以对它求导得到

$$\frac{\mathrm{d}(2\pi f_t t+0.82)}{\mathrm{d}t}=2\pi\left[\frac{3}{128}+\frac{3}{128}\left(\frac{2t}{1\,023}\right)\right]\equiv 2\pi\bar{f}_t,$$

其中 $\bar{f}_t$ 被称为是瞬间频率. 对于这个"线性鸣叫"时间序列，用哈尔、D(4)、C(6)和 LA(8)小波中的任意两个计算并绘出离散小波变换细节 $\mathcal{D}_k$，$k=1$，…，6. 用下面两个因素解释我们得到的细节序列：(a)序列$\{X_t\}$的瞬间频率；(b)你选择的小波的特性.

[4.20] 令$\mathcal{S}_{J_0,t}$是时间序列 $\mathbf{X}$ 的 J_0 层光滑$\mathcal{S}_{J_0}$的第 t 个元素. 证明

$$\frac{1}{N}\sum_{t=0}^{N-1}(\mathcal{S}_{J_0,t}-\bar{\mathcal{S}}_{J_0})^2=\frac{1}{N}\|\ \mathcal{S}_{J_0}\ \|^2-\bar{X}^2,$$

其中$\bar{\mathcal{S}}_{J_0}$和 $\bar{X}$ 分别是$\mathcal{S}_{J_0}$ 和 $\mathbf{X}$ 的样本均值(上式左面是$\mathcal{S}_{J_0}$的样本方差，右面和式(104c)中最后两项是一样的).

[4.21] 如果一个时间序列 $\mathbf{X}$ 的样本均值为零，那么这个时间序列的离散小波变换系数 $\mathbf{W}=\mathcal{W}\mathbf{X}$ 是否还为零?

[4.22] 验证式(59a)给出的 D(4)小波滤波器的平方增益函数是式(105b)给出的形式，其中 L 取为 4. 提示：利用滤波器级联的概念.

[4.23] 在围绕图 73 的讨论中，我们说的"3dB 降"点是对于哈尔和 D(4)小波标称通带为 $1/4\leqslant|f|\leqslant 1/2$ 的. 证明这个结论对 Daubechies 小波滤波器都是成立的.

[4.24] 式(110b)给出的滤波器$\{u_{2,l}\}$的相位函数为 $\theta_2(f)=-\pi f$，因此它是线性相位的. 能否通过前移整数 ν 将这个滤波器变成零相位的?

[4.25] 当 $N=2^J$ 时平移因子 $\nu_j^{(G)}$(式(114a))等于什么?

[4.26] 对于哈尔小波滤波器和 D(4)小波滤波器，结合下面给出的前移因子，绘出$\{h_l\}$的相位函数 $\theta^{(H)}(\cdot)$和$\{h_{j,l}\}$的相位函数 $\theta_j^{(H)}(\cdot)(j=2,3,4)$：

	$\{h_l\}$	$\{h_{2,l}\}$	$\{h_{3,l}\}$	$\{h_{4,l}\}$
哈尔	0	1	3	7
D(4)	2	5	11	23

(这些前移因子使得上述滤波器尽量接近于零相位). 要注意的是，如果 $z=x+\mathrm{i}y$，那么极坐标表示 $z=|z|\mathrm{e}^{\mathrm{i}\theta}$中的$\theta$ 可以用大多数的计算机语言计算，用反正切函数以 y/x 作参数，$\arctan y/x$ 就可以计算出这个角度θ，要注意 x 和 y 的符号.

[4.27] 对 coiflet 小波滤波器和尺度滤波器，$L=6$，12，18，24 和 30，建立一个类似表 137a 和 137b 的表，给出边界系数的个数.

第5章

极大重叠离散小波变换

5.0 引言

本章，我们描述称为极大重叠离散小波变换(Maximal Overlap Discrete Wavelet Transform，MODWT)的离散小波变换的一个修订版本. 这个名字来源于有关 Allan 方差的文献(参见 8.6 节)，其中基于哈尔小波滤波器的极大重叠离散小波变换从 20 世纪 70 年代就已经使用了(文献正文一个好的切入点是 Greenhall，1991；也可参见 Percival and Guttorp，1994). 本质上与极大重叠离散小波变换相同的变换(或有一些相似的变换)已经在一些小波文献中讨论过了，在这些文献中其名字为"非十进制离散小波变换"(Shensa，1992)，"移位不变离散小波变换"(Beylkin，1992；Lang 等，1995)，"小波框架"(Unser，1995)，"平移不变离散小波变换"(Coifman and Donoho，1995；Liang and Parks，1996；Del Marco and Weiss，1997)，"平稳离散小波变换"(Nason and Silverman，1995)，"时间不变离散小波变换"(Pesquet 等，1996)，"非十进制离散小波变换"(Bruce and Gao，1996a). 然而所有这些名字或多或少地对变换缺少精确的描述，我们优先选"极大重叠离散小波变换"，主要是因为它导致了很容易陈述("mod WT")的一个缩略词且具有"离散小波变换修改"的内涵.

与规范正交部分离散小波变换相比较，对于时间序列 $\boldsymbol{X}$ 的 J_0 层的极大重叠离散小波变换是产生维数为 N 的列向量 $\widetilde{\boldsymbol{W}}_1$，$\widetilde{\boldsymbol{W}}_2$，…，$\widetilde{\boldsymbol{W}}_{J_0}$ 和 $\widetilde{\boldsymbol{V}}_{J_0}$ 的一个高度冗余的非正交变换. 向量 $\widetilde{\boldsymbol{W}}_j$ 包含了与 $\boldsymbol{X}$ 在尺度 $\tau_j=2^{j-1}$ 上的变化相关的极大重叠离散小波变换的小波系数，而 $\widetilde{\boldsymbol{V}}_{J_0}$ 包含了与尺度 $\lambda_{J_0}=2^{J_0}$ 和更高尺度上的变化有关的极大重叠离散小波变换的尺度系数. 与离散小波变换一样，极大重叠离散小波变换是根据一种计算上有效的塔型算法来定义的. 区分极大重叠离散小波变换与离散小波变换有五个重要的特点：

[1] J_0 层的部分离散小波变换把抽样尺寸限制在 2^{J_0} 的整数倍上，而 J_0 层的极大重叠离散小波变换对任一抽样尺寸 N 都有很好的定义. 当 N 是 2^{J_0} 的整数倍时，部分离散小波变换可通过 $O(N)$ 次乘法计算出来，而相应的极大重叠离散小波变换需要 $O(N\log_2 N)$ 次乘法. 因此，对于使用极大重叠离散小波变换来说，有一个计算的代价，但它计算上的复杂性与广泛应用的快速傅里叶变换算法是相同的，因此通常是可以接受的.

[2]极大重叠离散小波变换可以用来形成一个多分辨分析(MRA)，这对离散小波变换来说也是正确的. 与通常的离散小波变换相比，这个多分辨分析的细节 $\widetilde{D}_j$ 和平滑 $\widetilde{S}_{J_0}$ 是用任意数值循环平移时间序列，并用相应的数值循环平移每个细节和平滑部分.

[3]与离散小波变换相比，极大重叠离散小波变换的细节和平滑部分是与零相位滤波器相关的，因此可以很容易地列出原时间序列在多分辨分析中的有意义的特征.

[4]极大重叠离散小波变换可用来形成基于小波和尺度系数的方差分析(ANOVA)，这对离散小波变换也是正确的；与离散小波变换相比，极大重叠离散小波变换的细节和平滑部分不能用来形成这样的分析.

[5]虽然一个时间序列和序列的一个循环平移有不同的基于离散小波变换的经验功率谱(如式(63a)和(63b)中表述的)，但相对应的基于极大重叠离散小波变换的谱是相同的. 事实上，我们仅将一个类似平移应用到原序列极大重叠离散小波变换的每个成分 $\widetilde{\mathbf{W}}_j$ 和 $\widetilde{\mathbf{V}}_{J_0}$ 上，就能得到一个循环平移时间序列的极大重叠离散小波变换；相反，循环序列的离散小波变换一般不能通过原序列离散小波变换成分 $\mathbf{W}_j$ 和 $\mathbf{V}_{J_0}$ 的任何循环平移得到.

我们在 5.1 节中说明了，当我们循环平移 $\mathbf{X}$ 时，就能产生通常的离散小波变换的问题，在 5.2 节中定义了基本的极大重叠离散小波变换的小波和尺度滤波器，并且用这些滤波器在 5.3 和 5.4 节启发并证明了对极大重叠离散小波变换的基本结果. 5.5 节我们为有效地计算极大重叠离散小波变换提出了一种塔式算法，紧接着在 5.6 节举出了一个极大重叠离散小波变换怎样处理一个循环平移时间序列的简单例子. 然后在 5.7 节到 5.10 节通过几个例子说明了极大重叠离散小波变换的用处. 我们以一些实际考虑因素的一个讨论(5.11 节)和本章里关键结果的总结(5.12 节)来结束.

5.1 循环平移对离散小波变换的影响

我们在极大重叠离散小波变换里的兴趣来源于下面这个事实，用离散小波变换对时间序列的分析主要依赖于我们在哪里“分组”序列，即把什么看作是分析的起点或起源. 在前文中我们已经注意到离散小波变换在离散哈尔小波经验功率谱对这样结果的灵敏度(参见式(63a)和(63b))，但是这里我们希望表示出起点是怎样影响离散小波变换和它相关的多分辨分析的(为比较，我们将回顾 5.6 节将提出的用极大重叠离散小波变换的例子).

图 161 的底行有时间序列 $\mathbf{X}$(第一和第三列)和用 5 个单位循环平移序列的某种形式，即 $\mathcal{T}^5\mathbf{X}$(第二和第四列)的完全相同的图. 时间序列有 $N=128$ 个抽样尺寸，且在 6 个值上是由零附近的一个“脉冲”构成的：

$$X_t=\begin{cases}\frac{1}{2}\cos\left(\frac{3\pi t}{16}+0.08\right), & t=40,\cdots,45;\\ 0, & t=0,\cdots,39 \text{ 和 } 46,\cdots,127.\end{cases}$$

图 161 的顶部五行表示出了 $\mathbf{X}$ 和 $\mathcal{T}^5\mathbf{X}$ 的 $J_0=4$ 层的部分离散小波变换(前两列)以及相关的多分辨分析(后两列). 离散小波变换都是建立在 LA(8) 小波滤波器的基础上的. 离散小波变换系数 $\mathbf{W}_j$ 和 $\mathbf{V}_1$ 像 4.8 节描绘的那样被绘制成图. 因

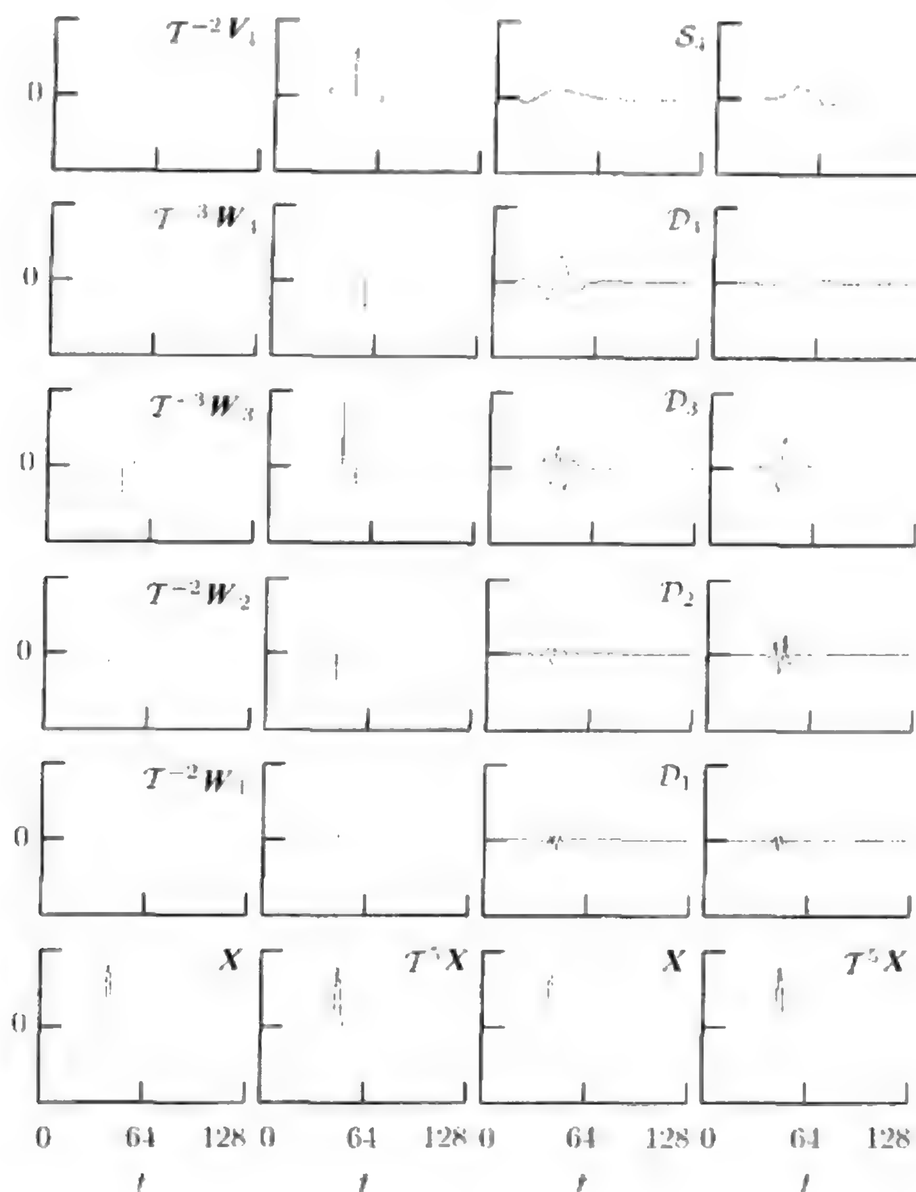


图 161 时间序列 $\mathbf{X}$ 及其循环平移 $\mathcal{T}^5\mathbf{X}$(图的底行)的 $J_0=4$ 层的 LA(8)离散小波变换(顶部五行，前面两列)和相应的多分辨分析(后面两列)．第 1 和第 2 列的比较表明循环平移一个时间序列能在其离散小波变换里产生实质变化；同样，第 3 和第 4 列表明了对相应的多分辨分析，这也是正确的．这个图应该与用极大重叠离散小波变换替代离散小波变换的图 181 进行比较

此，它们与基于 LA(8)滤波器的相位特性的原时间序列是一致的．注意，因为这些系数是作为零线的偏差而绘制成图的，故任何等于零的系数均以黑色线出现(由于脉冲和离散小波变换的局部性质，只有几个这样的系数)．尽管 $\mathbf{X}$ 和 $\mathcal{T}^5\mathbf{X}$ 中任一个的所有离散小波变换系数的平方和等于 $\|\mathbf{X}\|^2$，但前两列的比较表明，平移脉冲的时间起点已相当大地改变了离散小波变换．最后两列的一个检验表明，对多分辨分析来说，结论同样是正确的．虽然在每种情况里，细节和平滑部分的总和产生在底行的时间序列，对 $\mathbf{X}$ 和 $\mathcal{T}^5\mathbf{X}$ 的相应细节是完全不同的．

为了理解为什么平移时间起点能产生这些不同的离散小波变换，让我们集中精力在 $\mathbf{X}$ 和 $\mathcal{T}^5 X$ 的子向量 $\mathbf{W}_1$ 的元素 $W_{1,j}$($j=4$，…，7)上(这四个系数在图 161

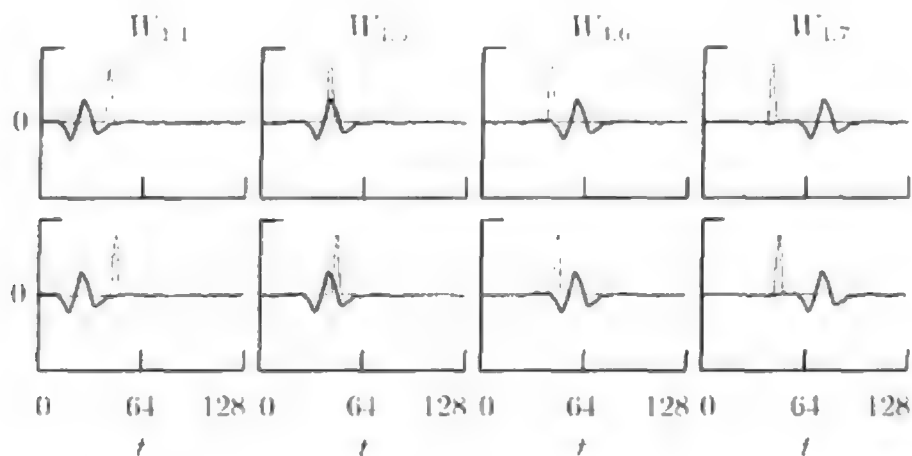


图 162　使用在 LA(8)离散小波变换中产生小波系数 $\boldsymbol{W}_{4,j}$，$j=4$，…，7（分别是图中两行从左到右的粗曲线）的 $J_0=4$ 层的基向量，连同一个“脉冲”时间序列 $\boldsymbol{X}$(图顶行的细曲线)和 $\mathcal{T}^5\boldsymbol{X}$(底行的细曲线)．图的基向量和时间序列之间的内积产生了图上方标记的小波系数

的第二行里的前两个图中表示出，恰恰是竖轴上零点标记符号的右边)．这些系数是由时间序列与一个向量的内积形成的，且这个向量的元素是在离散小波变换矩阵$\mathcal{W}$的行中．这四个行向量作为粗曲线在图 162 的顶行和底行绘制成图，连同(作为细曲线)在顶行的 $\boldsymbol{X}$ 和在底行的 $\mathcal{T}^5\boldsymbol{X}$ 绘制成图．如果我们取这个图顶行的任一个且计算两个绘出图的向量的内积，就得到 $\boldsymbol{X}$ 的一个系数 $W_{4,j}$；以一种类似的方式，底行产生 $\mathcal{T}^5\boldsymbol{X}$ 的相应系数．对于 $\boldsymbol{X}$ 或 $\mathcal{T}^5\boldsymbol{X}$，每个 $j=4$ 层的 LA(8)基向量是取两个加权平均之间的差分．第一个平均跨越 $\tau_3=2^3=8$ 的尺度，且被基向量的正的中心波瓣的位置所确定．第二个平均包含第一个平均之前和之后的点且被两个负的边波瓣位置所确定．如果我们考虑图 162 的第一行，就可以看到 $W_{4,5}$ 是很大的，因为 $\boldsymbol{X}$ 的脉冲被完全包含在 LA(8)基向量的中心波瓣内，然而两个负的边波瓣覆盖了是零的 $\boldsymbol{X}$ 的那一部分；另一方面，$W_{4,4}$，$W_{4,6}$ 和 $W_{4,7}$ 都接近于零，因为脉冲与中心波瓣或边波瓣是不一致的．图 162 的第二行表明，平移脉冲以得到 $\mathcal{T}^5\boldsymbol{X}$，改变了 $W_{4,5}$ 和 $W_{4,6}$ 的性质．对于 $W_{4,5}$，平移的脉冲覆盖了中心波瓣的部分和一个边波瓣的部分，因此产生的系数比在 $\boldsymbol{X}$ 这种情况里的更小；另一方面，对于 $W_{4,6}$，平移的脉冲现在与边波瓣中的一个相匹配，产生了一个负系数．

总之，虽然离散小波变换的小波系数可被解释为两个加权平均之间的差分，但这些平均产生的区间被严格地固定为一个先验的，且因此可能不与时间序列有意思的特性相一致．对时间序列起点上的变化，因此能产生了十分不同的结果，这是由于时间序列与离散小波变换预先定义的平均区间并列．和我们将要看到的一样，极大重叠离散小波变换是脱离对起点选择有利的结果的一种尝试，它本质上是通过包含所有可能的平均区间的设置得到的．

5.2 极大重叠离散小波变换的小波和尺度滤波器

为了很容易地将离散小波变换和极大重叠离散小波变换联系起来，我们在下面两节里将系统地阐述后者，通过$\widetilde{h}_l \equiv h_l/\sqrt{2}$定义极大重叠离散小波变换的小波滤波器$\{\widetilde{h}_l\}$和通过$\widetilde{g}_l \equiv g_l/\sqrt{2}$定义极大重叠离散小波变换的尺度滤波器$\{\widetilde{g}_l\}$是很方便的．因为$\{\widetilde{h}_l\}$仅是小波滤波器的一个简单改变尺度的形式，故很容易地从式(69a)、(69b)和(69c)得出，对所有非零整数n，

$$\sum_{l=0}^{L-1} \widetilde{h}_l = 0, \quad \sum_{l=0}^{L-1} \widetilde{h}_l^2 = \frac{1}{2} \quad 和 \quad \sum_{l=-\infty}^{\infty} \widetilde{h}_l \widetilde{h}_{l+2n} = 0. \tag{163a}$$

如果令$\widetilde{H}(\cdot)$和$\widetilde{\mathcal{H}}(\cdot)$分别表示对于$\{\widetilde{h}_l\}$的传递函数和平方增益函数，我们有$\widetilde{H}(f)=H(f)/\sqrt{2}$和$\widetilde{\mathcal{H}}(f)=\mathcal{H}(f)/2$，因此由式(69d)导出关系式

$$\widetilde{\mathcal{H}}(f)+\widetilde{\mathcal{H}}\left(f+\frac{1}{2}\right)=1 \quad 对所有的 f 成立. \tag{163b}$$

类似地，对极大重叠离散小波变换尺度滤波器，从练习[76b]得出，对所有的非零整数n，

$$\sum_{l=0}^{L-1} \widetilde{g}_l = 1, \quad \sum_{l=0}^{L-1} \widetilde{g}_l^2 = \frac{1}{2} \quad 和 \quad \sum_{l=-\infty}^{\infty} \widetilde{g}_l \widetilde{g}_{l+2n} = 0.$$

而由练习[77]得到，对所有的整数n，

$$\sum_{l=-\infty}^{\infty} \widetilde{g}_l \widetilde{h}_{l+2n} = 0.$$

如果令$\widetilde{G}(\cdot)$和$\widetilde{\mathcal{G}}(\cdot)$分别表示对$\{\widetilde{g}_l\}$的传递函数和平方增益函数，我们有$\widetilde{G}(f)=G(f)/\sqrt{2}$和$\widetilde{\mathcal{G}}(f)=\mathcal{G}(f)/2$，因此式(76b)告诉我们

$$\widetilde{\mathcal{G}}(f)+\widetilde{\mathcal{G}}\left(f+\frac{1}{2}\right)=1 \quad 和 \quad \widetilde{\mathcal{G}}(f)+\widetilde{\mathcal{H}}(f)=1 \quad 对所有 f 成立. \tag{163c}$$

还注意到，式(75a)和(75b)的正交镜像关系对极大重叠离散小波变换滤波器也是有效的：

$$\widetilde{g}_l = (-1)^{l+1}\widetilde{h}_{L-1-l} \quad 和 \quad \widetilde{h}_l = (-1)^l \widetilde{g}_{L-1-l}. \tag{163d}$$

用上述结果，我们可以将式(70b)和(77a)再表示为

$$\widetilde{W}_{1,t} \equiv \sum_{l=0}^{L-1} \widetilde{h}_l X_{t-l \bmod N} \quad 和 \quad \widetilde{V}_{1,t} \equiv \sum_{l=0}^{L-1} \widetilde{g}_l X_{t-l \bmod N}, \tag{163e}$$

$t=0,\cdots,N-1$．因此，序列$\{\widetilde{W}_{1,t}\}$和$\{\widetilde{V}_{1,t}\}$是分别用极大重叠离散小波变换的小波和尺度滤波器循环滤波$\{X_t\}$得到的．（在下一节里我们将会看到，这两个序列事实上构成了$J_0=1$层的极大重叠离散小波变换．）

5.3 极大重叠离散小波变换的基本概念

阐述极大重叠离散小波变换的动机在本质上是定义一个作用尽可能与离散小波变换一样的变换，但不会遭受离散小波变换对时间序列的起点选择的灵敏性. 这种灵敏性完全是由于在塔式算法的每一层上由小波和尺度滤波器下抽样(子抽样)输出的结果. 为得到对起点非灵敏的能力，我们需要在保存实施方差分析和多分辨分析的能力时，消除这种下抽样.

作为这个方面的第一步，注意，通常在离散小波变换塔式算法的第一步里被丢弃的滤波输出，可通过将离散小波变换塔式算法应用到循环平移向量$\mathcal{T}\mathbf{X}$而不是$\mathbf{X}$来得到. 这暗示了，当N是偶抽样尺寸时(事实上我们在以后会消除的一个假设)，为极大重叠离散小波变换定义塔式算法的第一步的下列程序. 想法是将通常离散小波变换塔式算法应用两次，一次对$\mathbf{X}$，一次对$\mathcal{T}\mathbf{X}$，此后将两组离散小波变换系数合并在一起. 第一次应用产生通常的离散小波变换，即

$$\begin{bmatrix}\mathbf{W}_1\\ \mathbf{V}_1\end{bmatrix}=\mathcal{P}_1\mathbf{X}=\begin{bmatrix}\mathcal{B}_1\\ \mathcal{A}_1\end{bmatrix}\mathbf{X},$$

其中(因为$\mathcal{W}_1=\mathcal{B}_1$，要求式(81b))

$$\mathcal{B}_1=\begin{bmatrix}h_1 & h_0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & h_3 & h_2\\ h_3 & h_2 & h_1 & h_0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & h_3 & h_2 & h_1 & h_0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & h_3 & h_2 & h_1 & h_0\end{bmatrix},$$

并且$\mathcal{A}_1$与将上式用g_l代替每个h_l有相同的结构(在这一章里，表示不同矩阵的元素时，我们清晰地特殊化了$L=4$和$N>L$，但我们的数学叙述对一般情况都是有效的). 考虑到式(70c)和(77b)，我们可以将$\mathbf{W}_1$和$\mathbf{V}_1$的元素表示为

$$\mathbf{W}_1=[2^{1/2}\widetilde{W}_{1,1},2^{1/2}\widetilde{W}_{1,3},\cdots,2^{1/2}\widetilde{W}_{1,N-1}]^{\mathrm{T}}$$

和

$$\mathbf{V}_1=[2^{1/2}\widetilde{V}_{1,1},2^{1/2}\widetilde{V}_{1,3},\cdots,2^{1/2}\widetilde{V}_{1,N-1}]^{\mathrm{T}};$$

即$\mathbf{W}_1$和$\mathbf{V}_1$包含所有长度为N的序列$\{2^{1/2}\widetilde{W}_{1,t}\}$和$\{2^{1/2}\widetilde{V}_{1,t}\}$的奇指标元素，其分别是用小波滤波器$\{h_l\}$和尺度滤波器$\{g_l\}$循环卷积时间序列$\mathbf{X}$形成的. 第二个应用是由循环平移向量$\mathcal{T}\mathbf{X}$代替$\mathbf{X}$得到

$$\begin{bmatrix}\mathbf{W}_{\mathcal{T},1}\\ \mathbf{V}_{\mathcal{T},1}\end{bmatrix}\equiv\mathcal{P}_1\mathcal{T}\mathbf{X}$$

来构成的. 等价地，如果我们定义

$$\mathcal{P}_{\mathcal{T},1}\equiv\mathcal{P}_1\mathcal{T}=\begin{bmatrix}\mathcal{B}_1\\ \mathcal{A}_1\end{bmatrix}\mathcal{T}=\begin{bmatrix}\mathcal{B}_1\mathcal{T}\\ \mathcal{A}_1\mathcal{T}\end{bmatrix}\equiv\begin{bmatrix}\mathcal{B}_{\mathcal{T},1}\\ \mathcal{A}_{\mathcal{T},1}\end{bmatrix},$$

我们可以写成

$$\begin{bmatrix} \mathbf{W}_{\mathcal{T},1} \\ \mathbf{V}_{\mathcal{T},1} \end{bmatrix} = \mathcal{P}_{\mathcal{T},1}\mathbf{X},$$

其中

$$\mathcal{B}_{\mathcal{T},1} = \begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & h_3 & h_2 & h_1 \\ h_2 & h_1 & h_0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & h_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & h_3 & h_2 & h_1 & h_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & h_3 & h_2 & h_1 & h_0 & 0 \end{bmatrix},$$

且$\mathcal{A}_{\mathcal{T},1}$有相似的结构而用 g_l代替每个 h_l. 比较$\mathcal{B}_1$ 和$\mathcal{B}_{\mathcal{T},1}$的内容说明，因为$\mathcal{B}_1\mathbf{X}$形成了滤波输出$\{2^{1/2}\widetilde{W}_{1,t}\}$的奇指标的值，所以$\mathcal{B}_{\mathcal{T},1}\mathbf{X}$ 必须形成偶指标的值. 因此，$\mathbf{W}_{\mathcal{T},1}$的元素是由

$$\mathbf{W}_{\mathcal{T},1} = [2^{1/2}\widetilde{W}_{1,0}, 2^{1/2}\widetilde{W}_{1,2}, \cdots, 2^{1/2}\widetilde{W}_{1,N-2}]^{\mathrm{T}}$$

给出的，并且通过一个类似的论断，$\mathbf{V}_{\mathcal{T},1}$的元素是由

$$\mathbf{V}_{\mathcal{T},1} = [2^{1/2}\widetilde{V}_{1,0}, 2^{1/2}\widetilde{V}_{1,2}, \cdots, 2^{1/2}\widetilde{V}_{1,N-2}]^{\mathrm{T}}$$

给出的. 下面，我们定义

$$\widetilde{\mathbf{W}}_1 \equiv [\widetilde{W}_{1,0}, \widetilde{W}_{1,1}, \widetilde{W}_{1,2}, \cdots, \widetilde{W}_{1,N-1}]^{\mathrm{T}}$$

和

$$\widetilde{\mathbf{V}}_1 \equiv [\widetilde{V}_{1,0}, \widetilde{V}_{1,1}, \widetilde{V}_{1,2}, \cdots, \widetilde{V}_{1,N-1}]^{\mathrm{T}};$$

即$\widetilde{\mathbf{W}}_1$ 是由改变尺度的$\mathbf{W}_1$ 和 $\mathbf{W}_{\mathcal{T},1}$的交叉元素形成的，且$\widetilde{\mathbf{V}}_1$ 是以一种类似的方式由$\mathbf{V}_1$ 和 $\mathbf{V}_{\mathcal{T},1}$构造. 注意$\widetilde{\mathbf{W}}_1$ 和$\widetilde{\mathbf{V}}_1$ 的元素确切的就是通过用极大重叠离散小波变换滤波器$\{\tilde{h}_l\}$和$\{\tilde{g}_l\}$得到的滤波器输出$\{\widetilde{W}_{1,t}\}$和$\{\widetilde{V}_{1,t}\}$——见式(163e). 让我们定义$\widetilde{\mathcal{B}}_1$ 是由交叉$\mathcal{B}_{\mathcal{T},1}$和$\mathcal{B}_1$ 的行用 $\tilde{h}_l$代替每个 h_l的 $N\times N$ 矩阵；即

$$\widetilde{\mathcal{B}}_1 \equiv \begin{bmatrix} \tilde{h}_0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 \\ \tilde{h}_1 & \tilde{h}_0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 & \tilde{h}_2 \\ \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 \end{bmatrix}. \quad (165a)$$

然后我们有$\widetilde{\mathbf{W}}_1 = \widetilde{\mathcal{B}}_1\mathbf{X}$. 对$\widetilde{\mathcal{A}}_1$ 有一个类似的定义，我们有$\widetilde{\mathbf{V}}_1 = \widetilde{\mathcal{A}}_1\mathbf{X}$. 最后，可以用

$$\begin{bmatrix} \widetilde{\mathbf{W}}_1 \\ \widetilde{\mathbf{V}}_1 \end{bmatrix} = \begin{bmatrix} \widetilde{\mathcal{B}}_1 \\ \widetilde{\mathcal{A}}_1 \end{bmatrix}\mathbf{X} = \widetilde{\mathcal{P}}_1\mathbf{X}, \quad \widetilde{\mathcal{P}}_1 \equiv \begin{bmatrix} \widetilde{\mathcal{B}}_1 \\ \widetilde{\mathcal{A}}_1 \end{bmatrix} \quad (165b)$$

表示极大重叠离散小波变换塔式算法的第一步.

因为$\mathcal{P}_1^{\mathrm{T}}\ \mathcal{P}_1=I_N$ 和$\mathcal{T}^{\mathrm{T}}\ \mathcal{T}=I_N$，由此得到

$$\mathcal{P}_{\mathcal{T},1}^{\mathrm{T}}\ \mathcal{P}_{\mathcal{T},1}=\mathcal{T}^{\mathrm{T}}\ \mathcal{P}_1^{\mathrm{T}}\ \mathcal{P}_1\ \mathcal{T}=I_N,$$

且因此 $\mathcal{P}_{\mathcal{T},1}$ 是一个规范正交矩阵. 因之我们也有

$$\|\mathbf{X}\|^2=\|\mathbf{W}_1\|^2+\|\mathbf{V}_1\|^2 \quad \text{和} \quad \|\mathbf{X}\|^2=\|\mathbf{W}_{\mathcal{T},1}\|^2+\|\mathbf{V}_{\mathcal{T},1}\|^2.$$

由于

$$\|\mathbf{W}_1\|^2+\|\mathbf{W}_{\mathcal{T},1}\|^2=2\|\widetilde{\mathbf{W}}_1\|^2 \quad \text{和} \quad \|\mathbf{V}_1\|^2+\|\mathbf{V}_{\mathcal{T},1}\|^2=2\|\widetilde{\mathbf{V}}_1\|^2,$$

由此得到

$$\|\mathbf{X}\|^2=\|\widetilde{\mathbf{W}}_1\|^2+\|\widetilde{\mathbf{V}}_1\|^2,$$

因此$\|\widetilde{\mathbf{W}}_1\|^2$ 和$\|\widetilde{\mathbf{V}}_1\|^2$分解$\|\mathbf{X}\|^2$，这是需要用极大重叠离散小波变换定义一个方差分析的基本结果.

因为我们有

$$\mathbf{X}=[\mathcal{B}_1^{\mathrm{T}},\mathcal{A}_1^{\mathrm{T}}]\begin{bmatrix}\mathbf{W}_1\\ \mathbf{V}_1\end{bmatrix} \quad \text{和} \quad \mathbf{X}=[\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}},\mathcal{A}_{\mathcal{T},1}^{\mathrm{T}}]\begin{bmatrix}\mathbf{W}_{\mathcal{T},1}\\ \mathbf{V}_{\mathcal{T},1}\end{bmatrix},$$

我们还可以通过平均上式两个右边项来复原 $\mathbf{X}$:

$$\begin{aligned}\mathbf{X}&=\frac{1}{2}[\mathcal{B}_1^{\mathrm{T}},\mathcal{A}_1^{\mathrm{T}}]\begin{bmatrix}\mathbf{W}_1\\ \mathbf{V}_1\end{bmatrix}+\frac{1}{2}[\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}},\mathcal{A}_{\mathcal{T},1}^{\mathrm{T}}]\begin{bmatrix}\mathbf{W}_{\mathcal{T},1}\\ \mathbf{V}_{\mathcal{T},1}\end{bmatrix}\\ &=\frac{1}{2}(\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1+\mathcal{A}_1^{\mathrm{T}}\mathbf{V}_1+\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}+\mathcal{A}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{V}_{\mathcal{T},1}).\end{aligned}$$

练习[166] 证明

$$\frac{1}{2}(\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1+\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1})=\widetilde{\mathcal{B}}_1^{\mathrm{T}}\ \widetilde{\mathbf{W}}_1$$

和

$$\frac{1}{2}(\mathcal{A}_1^{\mathrm{T}}\mathbf{V}_1+\mathcal{A}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{V}_{\mathcal{T},1})=\widetilde{\mathcal{A}}_1^{\mathrm{T}}\ \widetilde{\mathbf{V}}_1.$$ ◁

由上式，我们可以写成

$$\mathbf{X}=\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\mathbf{W}}_1+\widetilde{\mathcal{A}}_1^{\mathrm{T}}\widetilde{\mathbf{V}}_1\equiv\widetilde{\mathcal{D}}_1+\widetilde{\mathcal{S}}_1,$$

其中$\widetilde{\mathcal{D}}_1\equiv\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\mathbf{W}}_1$ 是第一层最大重叠细节，而$\widetilde{\mathcal{S}}_1\equiv\widetilde{\mathcal{A}}_1^{\mathrm{T}}\widetilde{\mathbf{V}}_1$ 是相应的平滑. 上面的公式是需要用极大重叠离散小波变换定义 $J_0=1$ 层的多分辨分析的基本结果.（练习[5.1]钻研了$\widetilde{\mathcal{D}}_1$ 和$\widetilde{\mathcal{S}}_1$ 的一个有趣的解释. 事实上，我们可以通过平均$\mathcal{D}_1$ 和$\mathcal{T}\mathcal{D}_{\mathcal{T},1}$来得到$\widetilde{\mathcal{D}}_1$，其中$\mathcal{D}_1$ 和$\mathcal{T}\mathcal{D}_{\mathcal{T},1}$分别是 $\mathbf{X}$ 和$\mathcal{T}\mathbf{X}$ 的第一层离散小波变换细节. 一个类似的论断对$\widetilde{\mathcal{S}}_1$ 也是有效的.）

重要的是，尽管我们对通常的离散小波变换有关系式$\|\mathbf{W}_1\|^2=\|\mathcal{D}_1\|^2$，但是如下面练习中指出的那样，一般在$\|\widetilde{\mathbf{W}}_1\|^2$ 和$\|\widetilde{\mathcal{D}}_1\|^2$ 之间没有等式.

练习[167a] 证明

$$\|\widetilde{\mathcal{D}}_1\|^2 = \frac{1}{2}(\|\widetilde{\mathbf{W}}_1\|^2 + \mathbf{W}_1^{\mathrm{T}} \mathcal{B}_1 \mathcal{B}_{\mathcal{T},1}^{\mathrm{T}} \mathbf{W}_{\mathcal{T},1}).$$ ◁

因此，尽管我们可用通常的离散小波变换作为$\|\mathbf{W}_1\|^2/N$或$\|\mathcal{D}_1\|^2/N$在尺度τ_1上定义离散小波经验功率谱，但当处理极大重叠离散小波变换时，我们仅能用$\|\widetilde{\mathbf{W}}_1\|^2/N$.（练习[5.7]表明，事实上$\|\widetilde{\mathcal{D}}_1\|^2\leqslant\|\widetilde{\mathbf{W}}_1\|^2$但这个练习最好是在学习了5.4节后再解决.）

上面极大重叠离散小波变换塔式算法的推导假设N是偶数，因此可将通常的离散小波变换塔式算法应用到$\mathbf{X}$和$\mathcal{T}\mathbf{X}$. 下面的练习说明这个算法的第一步事实上对所有的N都是有效的.

练习[167b] 对任意的抽样尺寸$N\geqslant L$，假设$\widetilde{\mathcal{B}}_1$是像式(165a)那样定义的，且类似地用$\tilde{g}_l$替代$\tilde{h}_l$来定义$\widetilde{\mathcal{A}}_1$. 假设$\widetilde{\mathbf{W}}_1$和$\widetilde{\mathbf{V}}_1$是由式(165b)定义的. 证明

$$\mathbf{X} = \widetilde{\mathcal{B}}_1^{\mathrm{T}} \widetilde{\mathbf{W}}_1 + \widetilde{\mathcal{A}}_1^{\mathrm{T}} \widetilde{\mathbf{V}}_1 \quad 和 \quad \|\mathbf{X}\|^2 = \|\widetilde{\mathbf{W}}_1\|^2 + \|\widetilde{\mathbf{V}}_1\|^2.$$

提示：证明$\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\mathcal{B}}_1+\widetilde{\mathcal{A}}_1^{\mathrm{T}}\widetilde{\mathcal{A}}_1=I_N$. ◁

事实上，和5.4节的讨论中所表明的一样，限制$N\geqslant L$可以放宽些.

用式(165b)所定义的$\widetilde{\mathcal{P}}_1$，我们可用$\mathbf{X}=\begin{bmatrix}\widetilde{\mathcal{B}}_1\\ \widetilde{\mathcal{A}}_1\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\widetilde{\mathbf{W}}_1\\ \widetilde{\mathbf{V}}_1\end{bmatrix}=\widetilde{\mathcal{P}}_1^{\mathrm{T}}\begin{bmatrix}\widetilde{\mathbf{W}}_1\\ \widetilde{\mathbf{V}}_1\end{bmatrix}$重新表示$\mathbf{X}=\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\mathbf{W}}_1+\widetilde{\mathcal{A}}_1^{\mathrm{T}}\widetilde{\mathbf{V}}_1$；进而因为$\begin{bmatrix}\widetilde{\mathbf{W}}_1\\ \widetilde{\mathbf{V}}_1\end{bmatrix}=\widetilde{\mathcal{P}}_1\mathbf{X}$，很明显对任何时间序列都有$\mathbf{X}=\widetilde{\mathcal{P}}_1^{\mathrm{T}}\widetilde{\mathcal{P}}_1\mathbf{X}$，因此$\widetilde{\mathcal{P}}_1^{\mathrm{T}}$可以认为是极大重叠离散小波变换$\widetilde{\mathcal{P}}_1$的逆. 因为$\widetilde{\mathcal{P}}_1$是一个$2N\times N$矩阵，事实上有无限种方法从其极大重叠离散小波变换系数$\widetilde{\mathbf{W}}_1$和$\widetilde{\mathbf{V}}_1$里复原$\mathbf{X}$，故$\widetilde{\mathcal{P}}_1^{\mathrm{T}}$决不是$\widetilde{\mathcal{P}}_1$的唯一的逆；然而，练习[5.2]将证明$\widetilde{\mathcal{P}}_1^{\mathrm{T}}$事实上是$\widetilde{\mathcal{P}}_1$的(唯一的)Moore-Penrose广义逆.

对于离散小波变换，我们在4.4节中讨论了细节$\mathcal{D}_1$可以用$\mathbf{W}_1$的上抽样形式通过互相关的周期化滤波器$\{h_l^{\circ}\}$构造出来. 这里，我们证明极大重叠离散小波变换细节$\widetilde{\mathcal{D}}_1$可以根据一个包含$\mathbf{X}$和一个零相位滤波器的滤波运算来解释. 令$\widetilde{\mathcal{D}}_{1,t}$是$\widetilde{\mathcal{D}}_1$的第$t$个元素，展开$\widetilde{\mathcal{D}}_1=\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\mathbf{W}}_1$得到

$$\widetilde{\mathcal{D}}_1 = \begin{bmatrix} \tilde{h}_0 & \tilde{h}_1 & \tilde{h}_2 & \tilde{h}_3 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \tilde{h}_0 & \tilde{h}_1 & \tilde{h}_2 & \tilde{h}_3 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \tilde{h}_0 & \tilde{h}_1 & \tilde{h}_2 & \tilde{h}_3 \\ \tilde{h}_3 & 0 & 0 & 0 & 0 & \cdots & 0 & \tilde{h}_0 & \tilde{h}_1 & \tilde{h}_2 \\ \tilde{h}_2 & \tilde{h}_3 & 0 & 0 & 0 & \cdots & 0 & 0 & \tilde{h}_0 & \tilde{h}_1 \\ \tilde{h}_1 & \tilde{h}_2 & \tilde{h}_3 & 0 & 0 & \cdots & 0 & 0 & 0 & \tilde{h}_0 \end{bmatrix} \begin{bmatrix} \widetilde{W}_{1,0} \\ \widetilde{W}_{1,1} \\ \widetilde{W}_{1,2} \\ \vdots \\ \widetilde{W}_{1,N-2} \\ \widetilde{W}_{1,N-1} \end{bmatrix},$$

由此我们有

$$\widetilde{\mathcal{D}}_{1,t} = \sum_{l=0}^{L-1} \tilde{h}_l \widetilde{W}_{1,t+l \bmod N} = \sum_{l=0}^{N-1} \tilde{h}_l^{\circ} \widetilde{W}_{1,t-l \bmod N}, \qquad t = 0,1,\cdots,N-1,$$

其中$\{h_l^\circ: l=0, \cdots, N-1\}$是$\{h_l: l=0, \cdots, L-1\}$以$N$为周期进行周期化所得的. 因此，细节序列可以通过$\{\tilde{h}_l^\circ\}$用循环互相关$\{\widetilde{W}_{1,t}\}$得到(参见式(37c)). 因为$\widetilde{H}(\cdot)$是$\{\tilde{h}_l\}$的传递函数，一个等价的论断是细节序列可以用离散傅里叶变换是$\left\{\widetilde{H}^*\left(\frac{k}{N}\right): k=0, \cdots, N-1\right\}$的一个滤波器循环滤波$\{\widetilde{W}_{1,t}\}$来得到. 因为$\{\widetilde{W}_{1,t}\}$是用$\{\tilde{h}_l^\circ\}$滤波$\{X_t\}$形成的，故细节序列可等价地用离散傅里叶变换是$\left\{\widetilde{H}^*\left(\frac{k}{N}\right)\right\}$的滤波器卷积$\{\tilde{h}_l^\circ\}$得到一个合成滤波器来滤波$\{\tilde{h}_l^\circ\}$形成. 现在合成循环滤波器的离散傅里叶变换是由单个滤波器离散傅里叶变换的积给出的，由此得到

$$\widetilde{H}^*\left(\frac{k}{N}\right)\widetilde{H}\left(\frac{k}{N}\right)=\left|\widetilde{H}\left(\frac{k}{N}\right)\right|^2=\widetilde{\mathcal{H}}\left(\frac{k}{N}\right).$$

因为$\widetilde{\mathcal{H}}(\cdot)$是实值的且是非负的，故对所有频率来说，其相位函数是零(如果我们像$f=0$这样定义它，这里$\widetilde{\mathcal{H}}(0)=0$)，这就建立了重要的结果：与$\widetilde{\mathcal{D}}_1$相关的合成滤波器有零相位. 用类似的论断，$\widetilde{S}_1$可用离散傅里叶变换是$\left\{\widetilde{G}^*\left(\frac{k}{N}\right)\right\}$的滤波器滤波$\{\widetilde{V}_{1,t}\}$来表示，这里$\widetilde{G}(\cdot)$是$\{\tilde{g}_l\}$的传递函数. 我们将应用到$\mathbf{X}$以得到$\widetilde{S}_1$的合成滤波器有一个由$\widetilde{\mathcal{G}}(f)\equiv|\widetilde{G}(f)|^2$定义的零相位传递函数的抽样所组成的离散傅里叶变换.

5.3 节的关键结论与定义

作为总结，N维向量$\widetilde{\mathbf{W}}_1$包含了$j=1$层的极大重叠离散小波变换小波系数. $\widetilde{\mathbf{W}}_1$的第t个元素是$\widetilde{W}_{1,t}$，$t=0, 1, \cdots, N-1$，且是用极大重叠离散小波变换小波滤波器$\{\tilde{h}_l\equiv h_l/\sqrt{2}\}$循环滤波时间序列$\{X_t\}$形成的(参见式(163e))；等价地，我们可以认为$\{\widetilde{W}_{1,t}\}$是用滤波器$\{\tilde{h}_l^\circ\}$循环滤波$\{X_t\}$的结果，其中$\{\tilde{h}_l^\circ\}$是$\{\tilde{h}_l\}$以N为周期进行周期化的结果. 同样地，极大重叠离散小波变换尺度系数$\widetilde{\mathbf{V}}_1$是用极大重叠离散小波变换尺度滤波器$\{\tilde{g}_l\equiv g_l/\sqrt{2}\}$滤波$\{X_t\}$得到的；等价地，我们可以认为$\{\widetilde{V}_{1,t}\}$是用$\{\tilde{g}_l^\circ\}$滤波$\{X_t\}$的结果，其中$\{\tilde{g}_l^\circ\}$是$\{\tilde{g}_l\}$周期化的形式. $N\times N$矩阵$\widetilde{\mathcal{B}}_1$的每一行都包含了周期化的滤波器$\{\tilde{h}_l^\circ\}$，行间的差别由循环平移给出；同样地，$\widetilde{\mathcal{A}}_1$的行元素是周期化的滤波器$\{\tilde{g}_l^\circ\}$. 有了这些定义，我们就有

$$\|\mathbf{X}\|^2=\|\widetilde{\mathbf{W}}_1\|^2+\|\widetilde{\mathbf{V}}_1\|^2 \quad 和 \quad \mathbf{X}=\widetilde{\mathcal{B}}_1^T\widetilde{\mathbf{W}}_1+\widetilde{\mathcal{A}}_1^T\widetilde{\mathbf{V}}_1\equiv\widetilde{\mathcal{D}}_1+\widetilde{\mathcal{S}}_1,$$

它们分别根据第1层极大重叠离散小波变换细节$\widetilde{\mathcal{D}}_1$和第1层极大重叠离散小波变换平滑$\widetilde{\mathcal{S}}_1$产生一个能量分解和一个加性分解. $\widetilde{\mathcal{D}}_1$的元素是用传递函数是$\widetilde{H}^*(\cdot)$的滤波器滤波$\widetilde{\mathbf{W}}_1$来得到的，其中$\widetilde{H}^*(\cdot)$是$\{\tilde{h}_l\}$的传递函数$\widetilde{H}(\cdot)$的复共轭；同样地，$\widetilde{\mathcal{S}}_1$的元素是用传递函数是$\widetilde{G}^*(\cdot)$的滤波器滤波$\widetilde{\mathbf{V}}_1$的结果，而$\widetilde{G}^*(\cdot)$是$\{\tilde{g}_l\}$的传递函数的复共轭. 借助极大重叠离散小波变换系数，描绘$\mathbf{X}$的分析和综合的流程图在图169中表示.

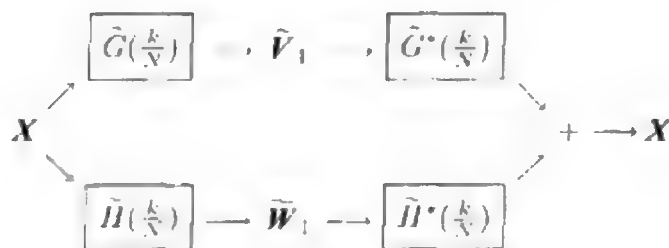


图 169 描述 $\boldsymbol{X}$ 到第一层极大重叠离散小波变换的小波和尺度系数$\widetilde{\boldsymbol{W}}_1$ 和$\widetilde{\boldsymbol{V}}_1$ 的分析流程图，紧接着就是由$\widetilde{\boldsymbol{W}}_1$ 和$\widetilde{\boldsymbol{V}}_1$ 对 $\boldsymbol{X}$ 的综合的流程图

5.4 第 j 层极大重叠离散小波变换系数的定义

对任意的抽样尺寸 N，我们现在定义第 j 层极大重叠离散小波变换小波和尺度系数为 N 维向量$\widetilde{\boldsymbol{W}}_j$ 和$\widetilde{\boldsymbol{V}}_j$，其元素分别是

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N} \quad 和 \quad \widetilde{V}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{g}_{j,l} X_{t-l \bmod N}, \tag{169a}$$

$t=0, \cdots, N-1$，其中 $\tilde{h}_{j,l} \equiv h_{j,l}/2^{j/2}$ 和 $\tilde{g}_{j,l} \equiv g_{j,l}/2^{j/2}$. 滤波器 $\{\tilde{h}_{j,l}\}$ 和 $\{\tilde{g}_{j,l}\}$ 被称为是第 j 层极大重叠离散小波变换小波和尺度滤波器. 注意，因为 $\{h_{j,l}\}$ 和 $\{g_{j,l}\}$ 均有宽度

$$L_j \equiv (2^j-1)(L-1)+1$$

(参见练习[96])，所以 $\{\tilde{h}_{j,l}\}$ 和 $\{\tilde{g}_{j,l}\}$ 的宽度也是一样的. 因为 $\{\tilde{h}_l\}$ 和 $\{\tilde{g}_l\}$ 的传递函数是由 $\widetilde{H}(f) \equiv H(f)/\sqrt{2}$ 和 $\widetilde{G}(f) \equiv G(f)/\sqrt{2}$ 给出的，故很明显地从式(96b)和(97)中看到，$\{\tilde{h}_{j,l}\}$ 和 $\{\tilde{g}_{j,l}\}$ 的传递函数分别是由

$$\widetilde{H}_j(f) \equiv \widetilde{H}(2^{j-1}f)\prod_{l=0}^{j-2} \widetilde{G}(2^l f) \quad 和 \quad \widetilde{G}_j(f) \equiv \prod_{l=0}^{j-1} \widetilde{G}(2^l f) \tag{169b}$$

给出. 对于 $j=1$，由上面的定义推出 $L_1 \equiv L$ 和 $\widetilde{G}_1(f) \equiv \widetilde{G}(f)$；我们还取

$$\widetilde{H}_1(f) \equiv \widetilde{H}(f), \quad \tilde{h}_{1,l} \equiv \tilde{h}_l \quad 和 \quad \tilde{g}_{1,l} \equiv \tilde{g}_l.$$

现在我们前面有两个任务. 第一，给定任意整数 $J_0 \geqslant 1$ 和任意抽样尺寸 N，我们需要证明，对于极大重叠离散小波变换，我们的定义产生能量分解

$$\|\boldsymbol{X}\|^2 = \sum_{j=1}^{J_0} \|\widetilde{\boldsymbol{W}}_j\|^2 + \|\widetilde{\boldsymbol{V}}_{J_0}\|^2 \tag{169c}$$

和加性分解

$$\boldsymbol{X} = \sum_{j=1}^{J_0} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_{J_0}, \tag{169d}$$

其中 $\widetilde{\mathcal{D}}_j$ 和 $\widetilde{\mathcal{S}}_{J_0}$ 以一种已经定义的方式分别通过$\widetilde{\boldsymbol{W}}_j$ 和$\widetilde{\boldsymbol{V}}_{J_0}$ 依赖于 $\boldsymbol{X}$. 注意上面的结果与 4.7 节对离散小波变换得出的结论是类似的(参见式(104b)和(104a)). 第

二，我们需要完成对极大重叠离散小波变换的塔式算法的描述，这在上节里仅对第一层描述了. 第一个任务是本节剩下部分的主题，而下节(5.5)将致力于极大重叠离散小波变换塔式算法的详述.

为了建立式(169c)，令$\{\tilde{h}^{\circ}_{j,l}\}$和$\{\tilde{g}^{\circ}_{j,l}\}$是$\{\tilde{h}_{j,l}\}$和$\{\tilde{g}_{j,l}\}$以 N 为周期进行周期化后得到的滤波器(参见 2.5 节). $\{\tilde{h}^{\circ}_{j,l}\}$和$\{\tilde{g}^{\circ}_{j,l}\}$的离散傅里叶变换因此分别由$\left\{\tilde{H}_j\left(\frac{k}{N}\right)\right\}$和$\left\{\tilde{G}_j\left(\frac{k}{N}\right)\right\}$给出. 因为用$\{\tilde{h}_{j,l}\}$循环滤波$\{X_t\}$与用$\{\tilde{h}^{\circ}_{j,l}\}$滤波是等价的，故我们可以将式(169a)重新表示为

$$\widetilde{W}_{j,t}=\sum_{l=0}^{N-1}\tilde{h}^{\circ}_{j,l}X_{t-l\bmod N}\ \text{和}\ \widetilde{V}_{j,t}=\sum_{l=0}^{N-1}\tilde{g}^{\circ}_{j,l}X_{t-l\bmod N},\tag{170a}$$

$t=0,\cdots,N-1$. 假设$\{\mathcal{X}_k\}$是$\{X_t\}$的离散傅里叶变换，我们有

$$\{\widetilde{W}_{j,t}\}\leftrightarrow\left\{\widetilde{H}_j\left(\frac{k}{N}\right)\mathcal{X}_k\right\}\quad\text{和}\quad\{\widetilde{V}_{j,t}\}\leftrightarrow\left\{\widetilde{G}_j\left(\frac{k}{N}\right)\mathcal{X}_k\right\}\tag{170b}$$

(参见式(37a)和(37b)). 帕塞瓦尔定理(式(36h))现在告诉我们

$$\|\widetilde{\mathbf{W}}_j\|^2=\frac{1}{N}\sum_{k=0}^{N-1}\left|\widetilde{H}_j\left(\frac{k}{N}\right)\right|^2|\mathcal{X}_k|^2\quad\text{和}\quad\|\widetilde{\mathbf{V}}_j\|^2=\frac{1}{N}\sum_{k=0}^{N-1}\left|\widetilde{G}_j\left(\frac{k}{N}\right)\right|^2|\mathcal{X}_k|^2,\tag{170c}$$

转而产生

$$\|\widetilde{\mathbf{W}}_j\|^2+\|\widetilde{\mathbf{V}}_j\|^2=\frac{1}{N}\sum_{k=0}^{N-1}|\mathcal{X}_k|^2\left(\left|\widetilde{H}_j\left(\frac{k}{N}\right)\right|^2+\left|\widetilde{G}_j\left(\frac{k}{N}\right)\right|^2\right).$$

对于任何 $j\geqslant 2$，可以使用式(169b)来简化上面括号中的项如下：

$$\begin{aligned}\left|\widetilde{H}_j\left(\frac{k}{N}\right)\right|^2+\left|\widetilde{G}_j\left(\frac{k}{N}\right)\right|^2&=\left|\widetilde{H}\left(2^{j-1}\frac{k}{N}\right)\right|^2\prod_{l=0}^{j-2}\left|\widetilde{G}\left(2^l\frac{k}{N}\right)\right|^2+\prod_{l=0}^{j-1}\left|\widetilde{G}\left(2^l\frac{k}{N}\right)\right|^2\\&=\left(\left|\widetilde{H}\left(2^{j-1}\frac{k}{N}\right)\right|^2+\left|\widetilde{G}\left(2^{j-1}\frac{k}{N}\right)\right|^2\right)\prod_{l=0}^{j-2}\left|\widetilde{G}\left(2^l\frac{k}{N}\right)\right|^2\\&=\left(\widetilde{\mathcal{H}}\left(2^{j-1}\frac{k}{N}\right)+\widetilde{\mathcal{G}}\left(2^{j-1}\frac{k}{N}\right)\right)\left|\widetilde{G}_{j-1}\left(\frac{k}{N}\right)\right|^2\\&=\left|\widetilde{G}_{j-1}\left(\frac{k}{N}\right)\right|^2,\end{aligned}\tag{170d}$$

其中我们已经用了对于所有的 f，有$\widetilde{\mathcal{H}}(f)+\widetilde{\mathcal{G}}(f)=1$ 的事实(即式(163c)). 现在，我们应用(170c)有

$$\|\widetilde{\mathbf{W}}_j\|^2+\|\widetilde{\mathbf{V}}_j\|^2=\frac{1}{N}\sum_{k=0}^{N-1}|\mathcal{X}_k|^2\left|\widetilde{G}_{j-1}\left(\frac{k}{N}\right)\right|^2=\|\widetilde{\mathbf{V}}_{j-1}\|^2.\tag{170e}$$

因为上式对 $j \geqslant 2$ 均成立，故通过归纳法的证明表明，对任一 $J_0 \geqslant 2$ 也有

$$\|\widetilde{\mathbf{V}}_1\|^2 = \sum_{j=2}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2.$$

如果我们能够证明 $\|\mathbf{X}\|^2 = \|\widetilde{\mathbf{W}}_1\|^2 + \|\widetilde{\mathbf{V}}_1\|^2$，则式(169c)的能量分解现在就能确立；事实上，练习[167b]表明当 $N \geqslant L$ 时，这是正确的，但这个限制在下面的练习中被去掉了.

练习[171a] 用一个与产生(170e)平行的论断证明

$$\|\mathbf{X}\|^2 = \|\widetilde{\mathbf{W}}_1\|^2 + \|\widetilde{\mathbf{V}}_1\|^2. \qquad \triangleleft$$

我们可用式(169c)的能量分解来提供基于下述 $\mathbf{X}$ 抽样方差分析的极大重叠离散小波变换. 因为$\hat{\sigma}_X^2 = \frac{1}{N}\|\mathbf{X}\|^2 - \overline{X}^2$（如式(48b)中所表述的），由此得到

$$\hat{\sigma}_X^2 = \frac{1}{N}\sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \frac{1}{N}\|\widetilde{\mathbf{V}}_{J_0}\|^2 - \overline{X}^2 \tag{171a}$$

（上式十分接近式(104c)表述的基于离散小波变换的方差分析）. 下面的练习证明了上式最后两项可认为是$\widetilde{\mathbf{V}}_{J_0}$的抽样方差.

练习[171b] 证明$\widetilde{\mathbf{V}}_{J_0}$的抽样均值等于 $\overline{X}$，因此$\widetilde{\mathbf{V}}_{J_0}$的抽样方差由$\frac{1}{N}\|\widetilde{\mathbf{V}}_{J_0}\|^2 - \overline{X}^2$ 给出. $\triangleleft$

作为准备建立基于式(169d)多分辨分析的极大重叠离散小波变换时，我们首先需要定义 $\widetilde{\mathcal{D}}_j$ 和$\widetilde{\mathcal{S}}_j$. 为这样做，先注意，用矩阵的记号，我们可以用

$$\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{W}}_j\mathbf{X} \quad 和 \quad \widetilde{\mathbf{V}}_j = \widetilde{\mathcal{V}}_j\mathbf{X} \tag{171b}$$

来表示从 $\mathbf{X}$ 到 $\widetilde{\mathbf{W}}_j$ 的和从 $\mathbf{X}$ 到 $\widetilde{\mathbf{V}}_j$ 的变换，其中 $N \times N$ 矩阵 $\widetilde{\mathcal{W}}_j$ 的每一行都具有由$\{\tilde{h}_{j,l}^{\circ}\}$指定的值，而 $\widetilde{\mathcal{V}}_j$ 具有由$\{\tilde{g}_{j,l}^{\circ}\}$指定的值；即鉴于式(170a)，我们一定有

$$\widetilde{\mathcal{W}}_j \equiv \begin{bmatrix} \tilde{h}_{j,0}^{\circ} & \tilde{h}_{j,N-1}^{\circ} & \tilde{h}_{j,N-2}^{\circ} & \tilde{h}_{j,N-3}^{\circ} & \cdots & \tilde{h}_{j,3}^{\circ} & \tilde{h}_{j,2}^{\circ} & \tilde{h}_{j,1}^{\circ} \\ \tilde{h}_{j,1}^{\circ} & \tilde{h}_{j,0}^{\circ} & \tilde{h}_{j,N-1}^{\circ} & \tilde{h}_{j,N-2}^{\circ} & \cdots & \tilde{h}_{j,4}^{\circ} & \tilde{h}_{j,3}^{\circ} & \tilde{h}_{j,2}^{\circ} \\ \tilde{h}_{j,2}^{\circ} & \tilde{h}_{j,1}^{\circ} & \tilde{h}_{j,0}^{\circ} & \tilde{h}_{j,N-1}^{\circ} & \cdots & \tilde{h}_{j,5}^{\circ} & \tilde{h}_{j,4}^{\circ} & \tilde{h}_{j,3}^{\circ} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \tilde{h}_{j,N-2}^{\circ} & \tilde{h}_{j,N-3}^{\circ} & \tilde{h}_{j,N-4}^{\circ} & \tilde{h}_{j,N-5}^{\circ} & \cdots & \tilde{h}_{j,1}^{\circ} & \tilde{h}_{j,0}^{\circ} & \tilde{h}_{j,N-1}^{\circ} \\ \tilde{h}_{j,N-1}^{\circ} & \tilde{h}_{j,N-2}^{\circ} & \tilde{h}_{j,N-3}^{\circ} & \tilde{h}_{j,N-4}^{\circ} & \cdots & \tilde{h}_{j,2}^{\circ} & \tilde{h}_{j,1}^{\circ} & \tilde{h}_{j,0}^{\circ} \end{bmatrix}, \tag{171c}$$

而$\widetilde{\mathcal{V}}_j$ 是由上式用$\{\tilde{g}_{j,l}^{\circ}\}$代替$\{\tilde{h}_{j,l}^{\circ}\}$来表示的. 与式(95c)类似，要求的定义是

$$\widetilde{\mathcal{D}}_j \equiv \widetilde{\mathcal{W}}_j^{\mathrm{T}}\widetilde{\mathbf{W}}_j \quad 和 \quad \widetilde{\mathcal{S}}_j \equiv \widetilde{\mathcal{V}}_j^{\mathrm{T}}\widetilde{\mathbf{V}}_j.$$

有了这些定义，现在我们能如下证明式(169d)的多分辨分析.

鉴于$\widetilde{\mathcal{W}}_j^T$和$\widetilde{\mathcal{V}}_j^T$的行的样子(参见式(171c))，我们可以将$\widetilde{\mathcal{D}}_j$和$\widetilde{\mathcal{S}}_j$的元素表示为

$$\widetilde{\mathcal{D}}_{j,t} = \sum_{l=0}^{N-1} \tilde{h}^{\circ}_{j,l} \widetilde{W}_{j,t+l \bmod N} \quad \text{和} \quad \widetilde{\mathcal{S}}_{j,t} = \sum_{l=0}^{N-1} \tilde{g}^{\circ}_{j,l} \widetilde{V}_{j,t+l \bmod N},$$

$t=0, \cdots, N-1$. 用与5.3节里建立$\widetilde{\mathcal{D}}_1$的元素是零相位滤波器输出的一个类似论断，我们知道$\{\widetilde{\mathcal{D}}_{j,t}\}$是用离散傅里叶变换是$\left\{\widetilde{H}_j^*\left(\frac{k}{N}\right)\right\}$的滤波器循环滤波$\{\widetilde{W}_{j,t}\}$的结果；同样地，$\{\widetilde{W}_{j,t}\}$是由离散傅里叶变换是$\left\{\widetilde{H}_j\left(\frac{k}{N}\right)\right\}$的滤波器$\{\tilde{h}^{\circ}_{j,l}\}$循环滤波$\{X_t\}$得到的. 令$\{\mathcal{X}_k\}$再一次表示$\{X_t\}$的离散傅里叶变换，由此得到

$$\{\widetilde{\mathcal{D}}_{j,t}\} \leftrightarrow \left\{\widetilde{H}_j^*\left(\frac{k}{N}\right)\widetilde{H}_j\left(\frac{k}{N}\right)\mathcal{X}_k\right\} = \left\{\left|\widetilde{H}_j\left(\frac{k}{N}\right)\right|^2 \mathcal{X}_k\right\}; \tag{172a}$$

同样地，我们有

$$\{\widetilde{\mathcal{S}}_{j,t}\} \leftrightarrow \left\{\left|\widetilde{G}_j\left(\frac{k}{N}\right)\right|^2 \mathcal{X}_k\right\}. \tag{172b}$$

因为$\{a_t\}\leftrightarrow\{A_k\}$和$\{b_t\}\leftrightarrow\{B_k\}$推出$\{a_t+b_t\}\leftrightarrow\{A_k+B_k\}$(一个简单的练习!)，我们有

$$\{\widetilde{\mathcal{D}}_{j,t}+\widetilde{\mathcal{S}}_{j,t}\} \leftrightarrow \left\{\left(\left|\widetilde{H}_j\left(\frac{k}{N}\right)\right|^2 + \left|\widetilde{G}_j\left(\frac{k}{N}\right)\right|^2\right)\mathcal{X}_k\right\}.$$

式(170d)表明当$j\geqslant 2$时$\left|\widetilde{H}_j\left(\frac{k}{N}\right)\right|^2+\left|\widetilde{G}_j\left(\frac{k}{N}\right)\right|^2=\left|\widetilde{G}_{j-1}\left(\frac{k}{N}\right)\right|^2$，因此上式变为

$$\{\widetilde{\mathcal{D}}_{j,t}+\widetilde{\mathcal{S}}_{j,t}\} \leftrightarrow \left\{\left|\widetilde{G}_{j-1}\left(\frac{k}{N}\right)\right|^2 \mathcal{X}_k\right\};$$

然而，等式(172b)表明也有$\{\widetilde{\mathcal{S}}_{j-1,t}\}\leftrightarrow\left\{\left|\widetilde{G}_{j-1}\left(\frac{k}{N}\right)\right|^2 \mathcal{X}_k\right\}$，故对所有的$t$，我们一定有$\widetilde{\mathcal{S}}_{j-1,t}=\widetilde{\mathcal{D}}_{j,t}+\widetilde{\mathcal{S}}_{j,t}$，这就建立了对所有的$j\geqslant 2$，

$$\widetilde{\mathcal{S}}_{j-1} = \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_j. \tag{172c}$$

一个简单的归纳论证表明对任意的$J_0\geqslant 2$有

$$\widetilde{\mathcal{S}}_1 = \sum_{j=2}^{J_0} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_{J_0}.$$

如果我们能证明$\mathbf{X}=\widetilde{\mathcal{S}}_1+\widetilde{\mathcal{D}}_1$，式(169d)的多分辨分析现在就建立了；虽然练习[167b]表明当$N\geqslant L$时，这是成立的，但在下面的练习中仍然可以消除这个限制.

练习[172]　用一个与导出式(172c)平行的论断证明

$$\mathbf{X} = \widetilde{\mathcal{S}}_1 + \widetilde{\mathcal{D}}_1. \qquad \triangleleft$$

5.4 节的关键结论与定义

对任意的抽样尺寸 N，极大重叠离散小波变换小波和尺度系数分别定义为

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N} \quad \text{和} \quad \widetilde{V}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{g}_{j,l} X_{t-l \bmod N},$$

$t=0, \cdots, N-1$，其中$\{\tilde{h}_{j,l}: l=0, \cdots, L_j-1\}$和$\{\tilde{g}_{j,l}: l=0, \cdots, L_j-1\}$分别是第 j 层极大重叠离散小波变换小波和尺度滤波器，且它们是通过 $\tilde{h}_{j,l} \equiv h_{j,l}/2^{j/2}$ 和$\tilde{g}_{j,l} \equiv g_{j,l}/2^{j/2}$根据第 j 层小波和尺度滤波器$\{h_{j,l}\}$和$\{g_{j,l}\}$来定义的(这里和以前一样 $L_j \equiv (2^j-1)(L-1)+1$)．用矩阵的记号，我们能将上式表示为

$$\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{W}}_j \mathbf{X} \quad \text{和} \quad \widetilde{\mathbf{V}}_j = \widetilde{\mathcal{V}}_j \mathbf{X},$$

其中$\widetilde{\mathcal{W}}_j$ 的行含有$\{\tilde{h}^{\circ}_{j,l}\}$的循环平移形式(即$\{\tilde{h}_{j,l}\}$以 N 为周期进行周期化)，而$\widetilde{\mathcal{V}}_j$的行包含$\{\tilde{g}^{\circ}_{j,l}\}$的循环平移形式．对于第 j 层极大重叠离散小波变换小波和尺度滤波器，其传递函数分别为

$$\widetilde{H}_j(f) \equiv \widetilde{H}(2^{j-1}f) \prod_{l=0}^{j-2} \widetilde{G}(2^l f)$$

和

$$\widetilde{G}_j(f) \equiv \prod_{l=0}^{j-1} \widetilde{G}(2^l f),$$

其中$\widetilde{H}(\cdot)$和$\widetilde{G}(\cdot)$是$\{\tilde{h}_l\} \equiv \{h_l/\sqrt{2}\}$和$\{\tilde{g}_l\} \equiv \{g_l/\sqrt{2}\}$的传递函数．定义$\widetilde{\mathbf{V}}_0 = \mathbf{X}$，我们对 $j=1, 2, \cdots$，有

$$\|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_j\|^2 = \|\widetilde{\mathbf{V}}_{j-1}\|^2,$$

这导出了对任意 $J_0 \geqslant 1$ 的极大重叠离散小波变换的能量分解

$$\|\mathbf{X}\|^2 = \sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2.$$

第 j 层极大重叠离散小波变换细节和平滑分别被定义为$\widetilde{\mathcal{D}}_j \equiv \widetilde{\mathcal{W}}_j^{\mathrm{T}} \widetilde{\mathbf{W}}_j$ 和$\widetilde{\mathcal{S}}_j \equiv \widetilde{\mathcal{V}}_j^{\mathrm{T}} \widetilde{\mathbf{V}}_j$，对任意的 $J_0 \geqslant 1$，借助于它，我们表示极大重叠离散小波变换加性分解

$$\mathbf{X} = \sum_{j=1}^{J_0} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_{J_0}.$$

5.4 节的评论与扩展

[1]如果抽样尺寸 N 是 2 的幂，时间序列 $\mathbf{X}$ 的离散小波变换就能因为关系式 $W_{j,t}=2^{j/2}\widetilde{W}_{j,2^j(t+1)-1}$ 从极大重叠离散小波变换里提取出来．对任意 m，循环平移序列 $\mathcal{T}^m\mathbf{X}$ 的离散小波变换能从极大重叠离散小波变换里提取出来也是正确的．Liang and Parks(1996)定义规范正交的"平移不变"离散小波变换时(类似概念由 Del Marco 和 Weiss 讨论，1997)，用了这个事实．他们考虑了 $\mathcal{T}^m\mathbf{X}(m=0,\cdots,N-1)$的离散小波变换，并且选取了某些能量最大化．这个过程用尽可能少的系数，集中到离散小波变换能量的一大部分的意义上，产生了最能匹配 $\mathbf{X}$ 的平移离散小波变换．这个变换实质上对循环平移是不灵敏的，因为它考虑了 $\mathbf{X}$ 的所有可能循环平移，且因此，如果我们用任意的 $\mathcal{T}^m\mathbf{X}$ 代替 $\mathbf{X}$ 作为我们的原时间序列，就能产生相同的结果．

因为离散小波变换需要 $O(N)$ 次乘法，所以使用没有理性的强制方法计算所有 N 的离散小波变换，以得到平移不变离散小波变换就需要计算 $O(N^2)$ 次乘法，但 Liang 和 Parks 指出极大重叠离散小波变换仅用 $O(N\log_2 N)$ 次乘法就能计算出所有需要的系数．此简化的基本原因是对 $\mathcal{T}^m\mathbf{X}$ 特定集进行离散小波变换，在特定尺度上有相同的系数．例如 $\mathcal{T}^m\mathbf{X}(m=0,2,\cdots,N-2)$ 和 $\mathcal{T}^m\mathbf{X}(m=1,3,\cdots,N-1)$有着相同的 $j=1$ 的系数；$\mathcal{T}^m\mathbf{X}(m=0,4,\cdots,N-4)$和 $\mathcal{T}^m\mathbf{X}(m=1,5,\cdots,N-3,m=2,6,\cdots,N-2$ 或 $m=3,7,\cdots,N-1)$一样有相同的 $j=2$ 的系数；等等．

[2]在 4.6 节中，我们讨论过将离散小波变换尺度系数$\{V_{j,t}\}$与尺度 $\lambda_j\equiv 2^j$ 联系起来是合理的，因为尺度滤波器$\{g_{j,l}\}$的自相关宽度总是等于λ_j：

$$\text{width}_a\{g_{j,l}\}\equiv\frac{\left(\sum_{l=-\infty}^{\infty}g_{j,l}\right)^2}{\sum_{l=-\infty}^{\infty}g_{j,l}^2}=\lambda_j$$

(这是式(103))．因为$\tilde{g}_{j,l}=g_{j,l}/2^{j/2}$，很容易可以看到也有 $\text{width}_a\{\tilde{g}_{j,l}\}=\lambda_j$，因为因子 $1/2^{j/2}$在上面的比率中消掉了．因此我们宣称极大重叠离散小波变换尺度系数$\{\widetilde{V}_{j,t}\}$是与尺度 λ_j 上的平均相关的，且用与离散小波变换例子中一样的原因，极大重叠离散小波变换小波系数$\{\widetilde{W}_{j,t}\}$是与尺度 $\tau_j\equiv 2^{j-1}$ 上的平均的变化相关的．

5.5　极大重叠离散小波变换的塔式算法

这里我们描述一种基于 $j-1$ 层的尺度系数$\widetilde{\mathbf{V}}_{j-1}$计算第 j 层极大重叠离散小波变换小波和尺度系数$\widetilde{\mathbf{W}}_j$ 和$\widetilde{\mathbf{V}}_j$ 的有效算法(这一算法的伪代码在评论和扩展的[1]中给出)．这一算法的关键是注意用来计算 $j-1$ 层和 j 层系数的滤波器之间的关系．式(170a)表明$\widetilde{\mathbf{W}}_j$、$\widetilde{\mathbf{V}}_j$ 和$\widetilde{\mathbf{V}}_{j-1}$ 的元素是分别用周期化的滤波器$\{\tilde{h}_{j,l}^{\circ}\}$、$\{\tilde{g}_{j,l}^{\circ}\}$和$\{\tilde{g}_{j-1,l}^{\circ}\}$循环滤波$\{X_t\}$得到的．这些滤波器的离散傅里叶变换分别是$\left\{\widetilde{H}_j\left(\frac{k}{N}\right)\right\}$、$\left\{\widetilde{G}_j\left(\frac{k}{N}\right)\right\}$和$\left\{\widetilde{G}_{j-1}\left(\frac{k}{N}\right)\right\}$．式(169b)的一个应用告诉我们这些离散傅里叶变换的前两个可以借助于第三个离散傅里叶变换和基本的小波和尺度滤波器$\{\tilde{h}_l\}$和$\{\tilde{g}_l\}$的传递函数$\widetilde{H}(\cdot)$和$\widetilde{G}(\cdot)$用下面的方式表示出来：

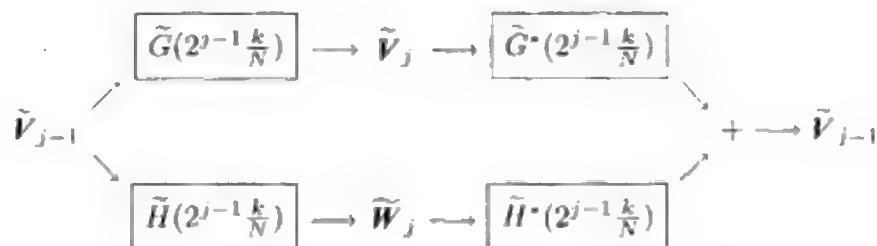


图 175　表示$\widetilde{\mathbf{V}}_{j-1}$到$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$的分析，接着是从$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$到$\widetilde{\mathbf{V}}_{j-1}$综合的流程图

$$\widetilde{H}_j\left(\frac{k}{N}\right)=\widetilde{G}_{j-1}\left(\frac{k}{N}\right)\widetilde{H}\left(2^{j-1}\frac{k}{N}\right)\quad 和\quad \widetilde{G}_j\left(\frac{k}{N}\right)=\widetilde{G}_{j-1}\left(\frac{k}{N}\right)\widetilde{G}\left(2^{j-1}\frac{k}{N}\right).$$

现在，上面等式的右边描述了由两个循环滤波器组成的滤波器级联．第一个滤波器—当被运用到 $\mathbf{X}$ 时——作为输出产生$\widetilde{\mathbf{V}}_{j-1}$．如果我们将这个输出应用到第二个循环滤波器，即离散傅里叶变换是$\left\{\widetilde{H}\left(2^{j-1}\frac{k}{N}\right)\right\}$和$\left\{\widetilde{G}\left(2^{j-1}\frac{k}{N}\right)\right\}$的滤波器，我们就能作为输出得到$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$．因此，通过用$\left\{\widetilde{H}\left(2^{j-1}\frac{k}{N}\right)\right\}$和$\left\{\widetilde{G}\left(2^{j-1}\frac{k}{N}\right)\right\}$描述的循环滤波器滤波$\widetilde{\mathbf{V}}_{j-1}$就能得到$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$.

为了精确地看到这两个滤波器采取什么样的形式，回忆练习[91]的结果：如果宽度为 L 的滤波器$\{\tilde{h}_l: l=0, \cdots, L-1\}$有传递函数$\widetilde{H}(\cdot)$，那么宽度为$2^{j-1}(L-1)+1$，脉冲响应序列为

$$\tilde{h}_0,\underbrace{0,\cdots,0}_{2^{j-1}-1个0},\tilde{h}_1,\underbrace{0,\cdots,0}_{2^{j-1}-1个0},\cdots,\tilde{h}_{L-2},\underbrace{0,\cdots,0}_{2^{j-1}-1个0},\tilde{h}_{L-1}\tag{175a}$$

的滤波器就有一个由$\{\widetilde{H}(2^{j-1}f)\}$定义的一个传递函数．因此用公式

$$\widetilde{W}_{j,t}=\sum_{l=0}^{L-1}\tilde{h}_l\widetilde{V}_{j-1,t-2^{j-1}l \bmod N},\quad t=0,1,\cdots,N-1\tag{175b}$$

可以从$\{\widetilde{V}_{j-1,t}\}$中得到元素$\{\widetilde{W}_{j,t}\}$；用类似的论断，我们也有

$$\widetilde{V}_{j,t}=\sum_{l=0}^{L-1}\tilde{g}_l\widetilde{V}_{j-1,t-2^{j-1}l \bmod N},\quad t=0,1,\cdots,N-1.\tag{175c}$$

这两个等式构成了极大重叠离散小波变换塔式算法．注意，如果我们定义$\widetilde{V}_{0,t}=X_t$，那上面的两个等式就产生第一层极大重叠离散小波变换小波和尺度系数$\widetilde{\mathbf{W}}_1$和$\widetilde{\mathbf{V}}_1$(参见式(163e)).

既然可以通过滤波$\widetilde{\mathbf{V}}_{j-1}$得到$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$，那如下面练习所描述的，从$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$重构$\widetilde{\mathbf{V}}_{j-1}$也是不足为奇的.

练习[175]　证明利用离散傅里叶变换是由$\left\{\widetilde{H}^*\left(2^{j-1}\frac{k}{N}\right)\right\}$和$\left\{\widetilde{G}^*\left(2^{j-1}\frac{k}{N}\right)\right\}$给出的滤波器，极大重叠离散小波逆变换可以通过下列等式描述的一种逆塔式算法来计算：

$$\widetilde{V}_{j-1,t}=\sum_{l=0}^{L-1}\tilde{h}_l\widetilde{W}_{j,t+2^{j-1}l \bmod N}+\sum_{l=0}^{L-1}\tilde{g}_l\widetilde{V}_{j,t+2^{j-1}l \bmod N},\quad t=0,1,\cdots,N-1.\tag{175d}$$

◁

图 175 描绘了极大重叠离散小波变换塔式算法和它的逆的流程图.

我们可以用 $N\times N$ 矩阵$\widetilde{\mathcal{B}}_j$和$\widetilde{\mathcal{A}}_j$表示从$\widetilde{\mathbf{V}}_{j-1}$到$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$的变换：

$$\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{B}}_j \widetilde{\mathbf{V}}_{j-1} \quad \text{和} \quad \widetilde{\mathbf{V}}_j = \widetilde{\mathcal{A}}_j \widetilde{\mathbf{V}}_{j-1}.$$

这些矩阵的内容从式(175b)和(175c)中是显而易见的. $\widetilde{\mathcal{B}}_j$的行是由式(175a)以 N 为周期进行周期化的上抽样小波滤波器组成的，其每一行由于向前或向后一个单位的循环平移而与它的邻行不同；同样地，$\widetilde{\mathcal{A}}_j$的行包含着由上抽样的尺度滤波器指定的值. 例如，对 $j=2$ 和 $N=12$，$L=4$，我们有

$$\widetilde{\mathcal{B}}_2 \equiv \begin{bmatrix} \tilde{h}_0 & 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 & \tilde{h}_1 & 0 \\ 0 & \tilde{h}_0 & 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 & \tilde{h}_1 \\ \tilde{h}_1 & 0 & \tilde{h}_0 & 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 & \tilde{h}_1 & 0 & \tilde{h}_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 & \tilde{h}_1 & 0 & \tilde{h}_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 & 0 & \tilde{h}_2 & 0 & \tilde{h}_1 & 0 & \tilde{h}_0 \end{bmatrix}, \tag{176}$$

而$\widetilde{\mathcal{A}}_2$是将上式的$\tilde{h}_l$用$\tilde{g}_l$替换得到的. 注意这个描述与 5.3 节$\widetilde{\mathcal{B}}_1$和$\widetilde{\mathcal{A}}_1$的公式是一致的. 有了$\widetilde{\mathcal{B}}_j$和$\widetilde{\mathcal{A}}_j$，我们就能用一种与离散小波变换类似的方式表示从$\widetilde{\mathbf{V}}_{j-1}$到$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$的极大重叠离散小波变换. 因此有

$$\begin{bmatrix} \widetilde{\mathbf{W}}_j \\ \widetilde{\mathbf{V}}_j \end{bmatrix} = \widetilde{\mathcal{P}}_j \widetilde{\mathbf{V}}_{j-1} = \begin{bmatrix} \widetilde{\mathcal{B}}_j \\ \widetilde{\mathcal{A}}_j \end{bmatrix} \widetilde{\mathbf{V}}_{j-1}, \quad \widetilde{\mathcal{P}}_j \equiv \begin{bmatrix} \widetilde{\mathcal{B}}_j \\ \widetilde{\mathcal{A}}_j \end{bmatrix},$$

这看起来与在所有分量顶上放上～的等式(94b)是相同的，也指明了极大重叠离散小波变换滤波器的用处. 进而，很明显地从式(175d)可以看到$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$到$\widetilde{\mathbf{V}}_{j-1}$的综合也可以根据$\widetilde{\mathcal{B}}_j$和$\widetilde{\mathcal{A}}_j$来表示：

$$\widetilde{\mathbf{V}}_{j-1} = \widetilde{\mathcal{B}}_j^{\mathrm{T}} \widetilde{\mathbf{W}}_j + \mathcal{A}_j^{\mathrm{T}} \widetilde{\mathbf{V}}_j.$$

如果我们回顾到$\widetilde{\mathbf{V}}_0 \equiv \mathbf{X}$，那上式直到 J_0 层的递推使用产生

$$\begin{aligned} \mathbf{X} = {} & \widetilde{\mathcal{B}}_1^{\mathrm{T}} \widetilde{\mathbf{W}}_1 + \widetilde{\mathcal{A}}_1^{\mathrm{T}} \widetilde{\mathcal{B}}_2^{\mathrm{T}} \widetilde{\mathbf{W}}_2 + \widetilde{\mathcal{A}}_1^{\mathrm{T}} \widetilde{\mathcal{A}}_2^{\mathrm{T}} \widetilde{\mathcal{B}}_3^{\mathrm{T}} \widetilde{\mathbf{W}}_3 + \cdots \\ & + \widetilde{\mathcal{A}}_1^{\mathrm{T}} \cdots \widetilde{\mathcal{A}}_{J_0-1}^{\mathrm{T}} \widetilde{\mathcal{B}}_{J_0}^{\mathrm{T}} \widetilde{\mathbf{W}}_{J_0} + \widetilde{\mathcal{A}}_1^{\mathrm{T}} \cdots \widetilde{\mathcal{A}}_{J_0-1}^{\mathrm{T}} \widetilde{\mathcal{A}}_{J_0}^{\mathrm{T}} \widetilde{\mathbf{V}}_{J_0}. \end{aligned}$$

我们可以用上式确定第 j 层极大重叠离散小波变换的细节$\widetilde{\mathcal{D}}_j$和第 J_0 层极大重叠离散小波变换的平滑$\widetilde{\mathcal{S}}_{J_0}$的表示式，即

$$\widetilde{\mathcal{D}}_j = \widetilde{\mathcal{A}}_1^{\mathrm{T}} \cdots \widetilde{\mathcal{A}}_{j-1}^{\mathrm{T}} \widetilde{\mathcal{B}}_j^{\mathrm{T}} \widetilde{\mathbf{W}}_j \quad \text{和} \quad \widetilde{\mathcal{S}}_{J_0} = \widetilde{\mathcal{A}}_1^{\mathrm{T}} \cdots \widetilde{\mathcal{A}}_{J_0-1}^{\mathrm{T}} \widetilde{\mathcal{A}}_{J_0}^{\mathrm{T}} \widetilde{\mathbf{V}}_{J_0}.$$

定义$\widetilde{\mathcal{D}}_j \equiv \widetilde{\mathcal{W}}_j^{\mathrm{T}} \widetilde{\mathbf{W}}_j$和$\widetilde{\mathcal{S}}_{J_0} \equiv \widetilde{\mathcal{V}}_{J_0}^{\mathrm{T}} \widetilde{\mathbf{V}}_{J_0}$，和上式的比较产生了根据$\mathcal{B}_j$和$\mathcal{A}_j$矩阵的$\widetilde{\mathcal{W}}_j$和$\widetilde{\mathcal{V}}_j$的表示式：

$$\widetilde{\mathcal{W}}_j = \widetilde{\mathcal{B}}_j \widetilde{\mathcal{A}}_{j-1} \cdots \widetilde{\mathcal{A}}_1 \quad \text{和} \quad \widetilde{\mathcal{V}}_j = \widetilde{\mathcal{A}}_j \widetilde{\mathcal{A}}_{j-1} \cdots \widetilde{\mathcal{A}}_1.$$

注意式(171b)说明通过用上边的 $N\times N$ 阶矩阵单独的左乘 $\mathbf{X}$ 就能得到$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$.

5.5 节的关键结论与定义

虽然用定义

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N} \quad 和 \quad \widetilde{V}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{g}_{j,l} X_{t-l \bmod N}, t = 0,1,\cdots,N-1$$

也可以从 $j-1$ 层的极大重叠离散小波变换的尺度系数经由下面的递推得到这些第 j 层的极大重叠离散小波变换系数(这里我们定义 $\widetilde{V}_{0,t} \equiv X_t$):

$$\widetilde{W}_{j,t} = \sum_{l=0}^{L-1} \tilde{h}_l \widetilde{V}_{j-1,t-2^{j-1}l \bmod N} \quad 和 \quad \widetilde{V}_{j,t} = \sum_{l=0}^{L-1} \tilde{g}_l \widetilde{V}_{j-1,t-2^{j-1}l \bmod N}.$$

但我们可以用矩阵的记号将上式描述为

$$\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{B}}_j \widetilde{\mathbf{V}}_{j-1} \quad 和 \quad \widetilde{\mathbf{V}}_j = \widetilde{\mathcal{A}}_j \widetilde{\mathbf{V}}_{j-1},$$

其中 $\widetilde{\mathcal{B}}_j$ 的行包含 $\{\tilde{h}_l\}$ 的循环平移形式，它已经被上抽样到宽度 $2^{j-1}(L-1)+1$，然后以 N 为周期被周期化. 基于 $\{\tilde{g}_l\}$，用类似的方式可以构造 $\widetilde{\mathcal{A}}_j$(这里，上抽样是通过在原始滤波器的 L 个值每两个间插上 $2^{j-1}-1$ 个 0 构成的). 我们可以通过

$$\widetilde{V}_{j-1,t} = \sum_{l=0}^{L-1} \tilde{h}_l \widetilde{W}_{j,t+2^{j-1}l \bmod N} + \sum_{l=0}^{L-1} \tilde{g}_l \widetilde{V}_{j,t+2^{j-1}l \bmod N}$$

从 $\{\widetilde{W}_{j,t}\}$ 和 $\{\widetilde{V}_{j,t}\}$ 重构 $\{\widetilde{V}_{j-1,t}\}$，这可以用矩阵的记号表示为 $\widetilde{\mathbf{V}}_{j-1} = \widetilde{\mathcal{B}}_j^T \widetilde{\mathbf{W}}_j + \widetilde{\mathcal{A}}_j^T \widetilde{\mathbf{V}}_j$. 上面的关系式允许我们将极大重叠离散小波变换的细节和平滑表示为

$$\widetilde{\mathcal{D}}_j = \widetilde{\mathcal{A}}_1^T \cdots \widetilde{\mathcal{A}}_{j-1}^T \widetilde{\mathcal{B}}_j^T \widetilde{\mathbf{W}}_j \quad 和 \quad \widetilde{\mathcal{S}}_{J_0} = \widetilde{\mathcal{A}}_1^T \cdots \widetilde{\mathcal{A}}_{J_0-1}^T \widetilde{\mathcal{A}}_{J_0}^T \widetilde{\mathbf{V}}_{J_0},$$

转而说明，矩阵 $\widetilde{\mathcal{W}}_j$ 和 $\widetilde{\mathcal{V}}_j$ 可以表示为

$$\widetilde{\mathcal{W}}_j = \widetilde{\mathcal{B}}_j \widetilde{\mathcal{A}}_{j-1} \cdots \widetilde{\mathcal{A}}_1 \quad 和 \quad \widetilde{\mathcal{V}}_j = \widetilde{\mathcal{A}}_j \widetilde{\mathcal{A}}_{j-1} \cdots \widetilde{\mathcal{A}}_1$$

(这些矩阵产生 $\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{W}}_j \mathbf{X}$ 和 $\widetilde{\mathbf{V}}_j = \widetilde{\mathcal{V}}_j \mathbf{X}$).

5.5 节的评论与扩展

[1]这里我们给出用来计算极大重叠离散小波变换及其逆的伪代码. 对 $t=0$, …, $N-1$，定义 $\widetilde{V}_{0,t}=X_t$，其中 N 为正整数. 设 $\{\tilde{h}_l: l=0, \cdots, L-1\}$ 是宽度 L 为偶数的一个极大重叠离散小波变换小波滤波器. 设 $\{\tilde{g}_l\}$ 是由式(163d)从 $\{\tilde{h}_l\}$ 构造出的相应的极大重叠离散小波变换尺度滤波器. 将 $\widetilde{\mathbf{W}}_j$ 和 $\widetilde{\mathbf{V}}_{J_0}$ 的元素表示为 $\widetilde{W}_{j,t}$ 和 $\widetilde{V}_{J_0,t}$，$t=0$, …, $N-1$. 给出 $\widetilde{\mathbf{V}}_{j-1}$，下面的伪代码计算出 $\widetilde{\mathbf{W}}_j$ 和 $\widetilde{\mathbf{V}}_j$(这是基于(175b)和(175c)):

对 $t=0$, …, $N-1$，进行下面的外循环:

置 $k := t$.

置 $\widetilde{W}_{j,t} := \tilde{h}_0 \widetilde{V}_{j-1,k}$.

置 $\widetilde{V}_{j,t}$ t := $\tilde{g}_0 \widetilde{V}_{j-1,k}$.

对 $n=1, \cdots, L-1$，进行下面的内循环：

k 减少 2^{j-1}.

如果 $k<0$，置 $k:=k \bmod N$.

$\widetilde{W}_{j,t}$ 增加 $\tilde{h}_n \widetilde{V}_{j-1,k}$.

$\widetilde{V}_{j,t}$ 增加 $\tilde{g}_n \widetilde{V}_{j-1,k}$.

结束内循环.

结束外循环.

注意，如果 j 满足 $2^{j-1} \leqslant N$（在实际应用中通常会遇到），那么以上代码中的单个的“$k \bmod N$”就能用“ $k+N$ ”代替．通过先用 $j=1$ 时上面的伪代码（即由 $\widetilde{V}_0=X$ 开始）且对 $j=2, \cdots, J_0$，我们能得到极大重叠离散小波变换的组成向量，即 $\widetilde{W}_1, \cdots, \widetilde{W}_{J_0}$ 和 $\widetilde{V}_{J_0}$，而且还有中间向量 $\widetilde{V}_1, \cdots, \widetilde{V}_{J_0-1}$.

极大重叠离散小波逆变换的任务就是给出向量 $\widetilde{W}_1, \cdots, \widetilde{W}_{J_0}$ 和 $\widetilde{V}_{J_0}$ 计算 X. 给出 $\widetilde{W}_j$ 和 $\widetilde{V}_j$，下面的伪代码用来计算 $\widetilde{V}_{j-1}$（这是基于式(175d))：

对 $t=0, \cdots, N-1$，进行下面的外循环：

置 $k:=t$.

置 $\widetilde{V}_{j-1,t}:=\tilde{h}_0 \widetilde{W}_{j,k}+\tilde{g}_0 \widetilde{V}_{j,k}$.

对 $n=1, \cdots, L-1$，进行下面的内循环：

k 减少 2^{j-1}.

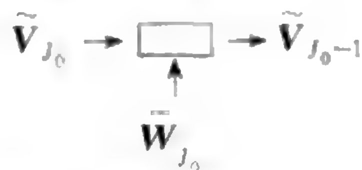
如果 $k \geqslant N$，置 $k:=k \bmod N$.

$\widetilde{V}_{j-1,t}$ 增加 $\tilde{h}_n \widetilde{W}_{j,k}+\tilde{g}_n \widetilde{V}_{j,k}$.

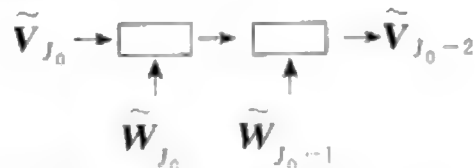
结束内循环.

结束外循环.

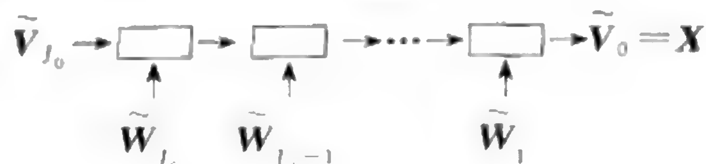
再次，如果 j 满足 $2^{j-1} \leqslant N$，那么上面的伪代码中单个的“$k \bmod N$”就能用“ $k-N$”来代替．通过先用 $\widetilde{W}_{J_0}$ 和 $\widetilde{V}_{J_0}$ 的伪代码，我们得到 $\widetilde{V}_{J_0-1}$．我们可以将 $\widetilde{W}_{J_0}$ 和 $\widetilde{V}_{J_0}$ 放进箱子里，出来的是 $\widetilde{V}_{J_0-1}$，来描绘这个程序：



接下来用 $\widetilde{W}_{J_0-1}$ 和 $\widetilde{V}_{J_0-1}$ 的伪代码产生 $\widetilde{V}_{J_0-2}$：

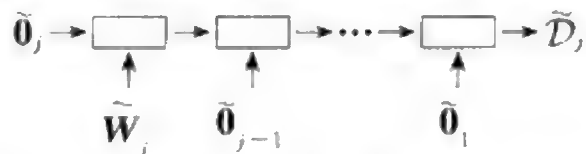


经过 J_0-2 次伪代码的反复应用，我们得到 $\widetilde{V}_0=X$：



($\widetilde{\mathbf{V}}_{J_0-3}$，…，$\widetilde{\mathbf{V}}_1$ 是作为中间量计算而产生的.)

第 j 层细节 $\widetilde{\mathcal{D}}_j$ 是用 $\tilde{\mathbf{0}}_1$，…，$\tilde{\mathbf{0}}_{j-1}$，$\widetilde{\mathbf{W}}_j$ 和 $\tilde{\mathbf{0}}_j$ 的极大重叠离散小波逆变换得到的，其中每个 $\tilde{\mathbf{0}}_k(k-1，…，j)$均是有 N 个零的向量. 这个细节在上面伪代码里用$\widetilde{\mathbf{W}}_j$ 和 $\tilde{\mathbf{0}}_j$ 代替$\widetilde{\mathbf{V}}_j$ 出发，能通过逆推应用 j，$j-1$，…，1 的逆极大重叠离散小波变换算法得到. 在 j 次迭代的末尾，我们得到了希望的 $\widetilde{\mathcal{D}}_j$：



同样地，平滑$\widetilde{\mathcal{S}}_{J_0}$ 可以通过应用极大重叠离散小波逆变换到 $\tilde{\mathbf{0}}_1$，…，$\tilde{\mathbf{0}}_{J_0}$ 和$\widetilde{\mathbf{V}}_{J_0}$ 得到：



[2]极大重叠离散小波变换细节和平滑与零相位滤波器相关的事实是一个以任何循环滤波 $\{a_l^\circ\} \leftrightarrow \left\{A\left(\frac{k}{N}\right)\right\}$ 为基础得到的一个零相位输出的著名"技巧"的例子. 技巧(例如，在 Hamming，1989，第 252 页描述的)就是用 $\{a_l^\circ\}$ 滤波时间序列 $\mathbf{X}$；反转从这个滤波器中的输出；然后用 $\{a_l^\circ\}$ 滤波被反转过的输出再次得到一个输出——称再反转得到的为 $\mathbf{Y}$——与零相位相关. 因为滤波反转过的输出，接着再被反转，是与原输出的互相关作用相同的事情，这个过程等价于下面的流程图：

$$\mathbf{X} \longrightarrow \boxed{A\left(\frac{k}{N}\right)} \longrightarrow \boxed{A^*\left(\frac{k}{N}\right)} \longrightarrow \mathbf{Y};$$

即二级级联的等价循环滤波是具有 $\left\{\left|A\left(\frac{k}{N}\right)\right|^2\right\}$ 的频域表示，其必然有零相位.

5.6 "脉冲"时间序列的极大重叠离散小波变换分析

现在让我们回顾一下 5.1 节讨论过的例子，在那里我们看到了一个"脉冲"时间序列 $\mathbf{X}$ 和它的循环平移 $\mathcal{T}^5\mathbf{X}$ 的离散小波变换分析. 图 181 底部的六行表明了与图 161 中相同的量，但现在我们用极大重叠离散小波变换代替离散小波变换. 让我们先考虑 $\mathbf{X}$ 的极大重叠离散小波变换系数(第 1 列，中间五行). 与离散小波变换相比较，每个极大重叠离散小波变换系数向量$\widetilde{\mathbf{W}}_j$ 和$\widetilde{\mathbf{V}}_4$ 都与 $\mathbf{X}$ 有相同数目的元素. 如第一列里标识符指示的那样，我们向前平移每个系数后(即向左边循环平移)，再绘制成图(参见式(52b)；例如，$\mathcal{T}^{-53}\widetilde{\mathbf{W}}_4$ 的第一个和最后一个值分别是 $\widetilde{W}_{4,53}$ 和$\widetilde{W}_{1,52}$). 因为在这个例子里，我们用 LA(8)小波滤波器，故这些循环平移前移的极大重叠离散小波变换系数向量使他们近似是零相位滤波器的输出. 对于小波和尺度系数，前移的特定值分别是由式(114b)中的 $\nu_j^{(H)}$ 和式((114a)中的 $\nu_4^{(G)}$ 的绝对值所决定的(在这些式中，对 LA(8)滤波器，$\nu=-3$). 对平移序列 $\mathcal{T}^5\mathbf{X}$，相应的

极大重叠离散小波变换系数向量在第 2 列中表示出．严格比较第 1 列和第 2 列说明，平移 $\mathbf{X}$ 有着用类似数值平移每个极大重叠离散小波变换系数向量的效果；即如果$\widetilde{\mathbf{W}}_j$ 和$\widetilde{\mathbf{V}}_{J_0}$ 构成了 $\mathbf{X}$ 的极大重叠离散小波变换，那么$\mathcal{T}^m\widetilde{\mathbf{W}}_j$ 和$\mathcal{T}^m\widetilde{\mathbf{V}}_{J_0}$ 构成了 $\mathcal{T}^m\mathbf{X}$ 的极大重叠离散小波变换．与离散小波变换相比较而言，循环平移 $\mathbf{X}$ 以一种直观合理的方法修改了极大重叠离散小波变换系数．还注意到，最大的极大重叠离散小波变换系数出现在$\widetilde{\mathbf{W}}_3$中，即在尺度 $\tau_3=4$ 的变化上．假定脉冲仅在六个值上是非零的且有 4.3 的有效宽度，这看起来是合理的(用与式(103)类似的一个有限抽样尺寸提供的一种度量)．

让我们比较 $\mathbf{X}$ 和$\mathcal{T}^5\mathbf{X}$ 的多分辨分析——这些分别在图 181 的第 3 和第 4 列表示(中间五行)．对于两个序列的极大重叠离散小波变换细节和平滑再一次用一个简单合理的方式连接起来：如果$\widetilde{\mathcal{D}}_j$ 和$\widetilde{\mathcal{S}}_{J_0}$ 形成了 $\mathbf{X}$ 的多分辨分析，那么$\mathcal{T}^m\widetilde{\mathcal{D}}_j$ 和$\mathcal{T}^m\widetilde{\mathcal{S}}_{J_0}$ 就形成了$\mathcal{T}^m\mathbf{X}$ 的多分辨分析．从图 161 很明显地可以看到，这个简单的关系式对离散小波变换来说是不成立的．

最后，我们比较一下对脉冲序列的极大重叠离散小波变换系数和多分辨分析．通过观察图 181 的第一列和第三列里的特定行，可以看到，向量$\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j$ 和$\widetilde{\mathcal{D}}_j$ 在形式上是十分类似的，$\mathcal{T}^{-|\nu_4^{(G)}|}\widetilde{\mathbf{V}}_4$ 和$\widetilde{\mathcal{S}}_4$ 也是一样．对比之下，图 161 的相应部分表明离散小波变换细节$\mathcal{D}_j$ 看起来与$\widetilde{\mathbf{W}}_j$ 完全不同(在一定程度上因为前者比后者多出 2^j 个元素)．为了弄清楚$\widetilde{\mathbf{W}}_j$ 和$\widetilde{\mathcal{D}}_j$ 为何看起来类似，回忆两个向量都是作为循环滤波 $\mathbf{X}$ 的输出而产生的．对$\widetilde{\mathbf{W}}_j$ 来说，循环滤波器有$\left\{H_j\left(\frac{k}{N}\right)\right\}$的频域表示．因为 $\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j$ 近似地是零相位滤波器的输出，故它的相关循环滤波器近似是$\left\{\left|H_j\left(\frac{k}{N}\right)\right|\right\}$．另一方面，从式(172a)可以知道，产生$\widetilde{\mathcal{D}}_j$ 的循环滤波器是由$\left\{\left|H_j\left(\frac{k}{N}\right)\right|^2\right\}$给出的，故产生$\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j$ 和$\widetilde{\mathcal{D}}_j$ 的循环滤波器就有着类似的相位和增益特性($H_j(\cdot)$ 近似上是一个带通滤波器的事实意味着 $|H_j(\cdot)|^2$ 也必须是——可以证明后者在某些度量标准下是一个更好的逼近)．在离散小波变换情形下，我们就没有这样一个紧密的对应，主要是因为滤波现在是通过先下抽样形成$\mathbf{W}_j$ 然后再上抽样形成 $\mathcal{D}_j$ 得到的．

虽然极大重叠离散小波变换细节和平移的小波系数是从有类似相位和增益特性的滤波器中输出的，但他们能被充分利用的方式是十分不同的．让我们用脉冲序列 $\mathbf{X}$ 和$\mathcal{T}^5\mathbf{X}$ 来说明．首先，虽然式(169c)表明我们可以用极大重叠离散小波变换系数分解在 $\mathbf{X}$ 或$\mathcal{T}^5\mathbf{X}$ 里的能量，但对细节和光滑这是不正确的．为证明它，再次基于 LA(8)小波滤波器计算 $J_0=7$ 层的极大重叠离散小波变换，并构造离散小波经验功率谱：

$$P_{\widetilde{W}}(\tau_j)\equiv\frac{1}{N}\|\widetilde{\mathbf{W}}_j\|^2=\frac{1}{N}\|\mathcal{T}^5\widetilde{\mathbf{W}}_j\|^2,\quad j=1,\cdots,7.$$

由于脉冲序列有 $N=128=2^7$ 个点，我们能用练习[5.9a]断言这些谱能为 $\mathbf{X}$ 或$\mathcal{T}^5\mathbf{X}$ 提供抽样方差的一个分析：

$$\hat{\sigma}_X^2=\sum_{j=1}^{7}P_{\widetilde{W}}(\tau_j)$$

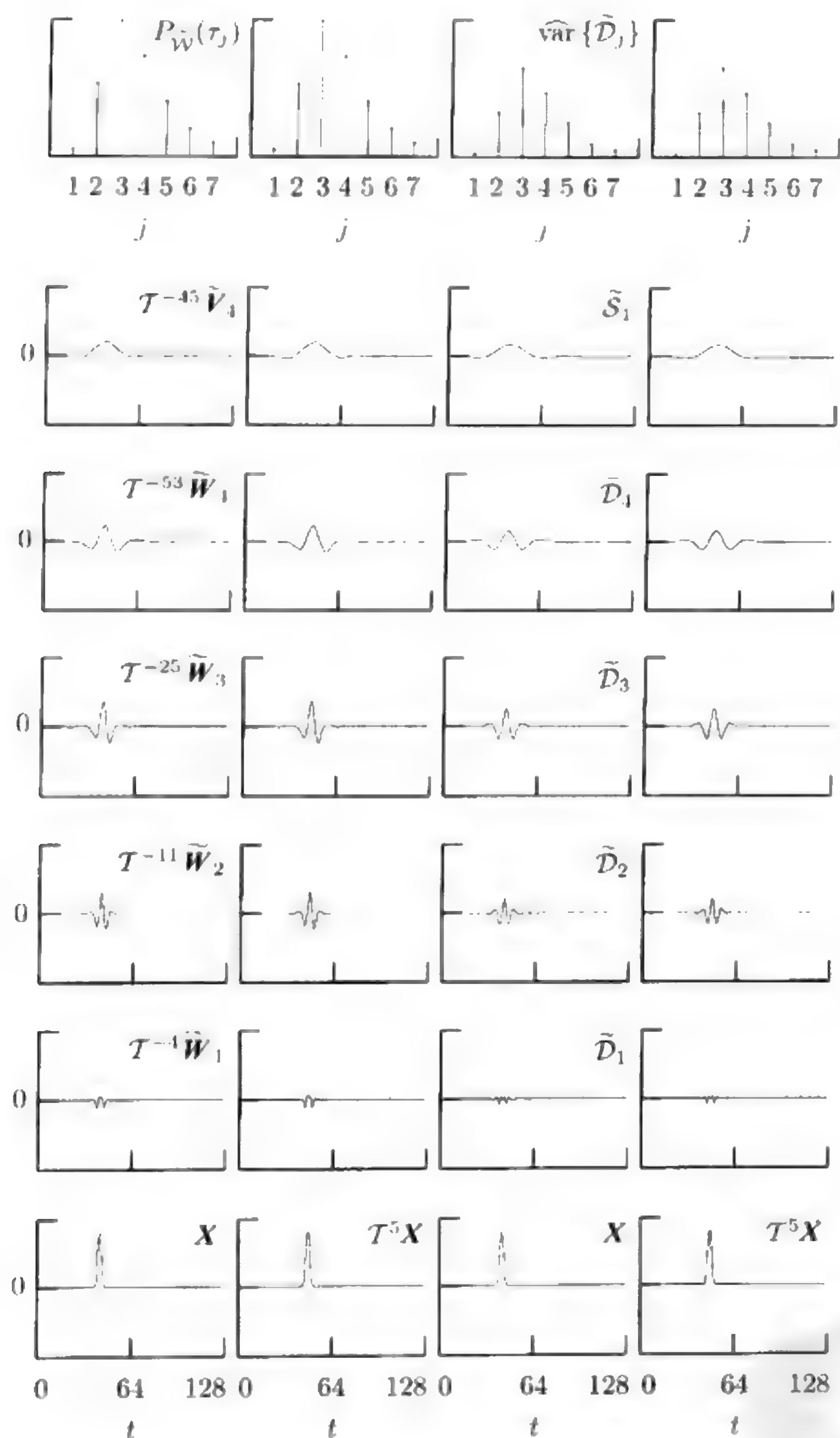


图 181 用极大重叠离散小波变换代替图 161 离散小波变换作为下面六行. 顶行表示离散小波经验功率谱(前两列)和一个基于 $j=1, \cdots, 7$ 层的极大重叠离散小波变换细节相应的量(后两列)——详情参见正文

(根据极大重叠离散小波变换系数对$\hat{\sigma}_X^2$ 的分析将在第 8 章里进一步研究). 图 181 第一行的前两个图表示出了 $\boldsymbol{X}$ 和$\mathcal{T}^5\boldsymbol{X}$ 的谱, 并且看起来是一样的——循环平移一个序列不改变它基于极大重叠离散小波变换上的功率谱(式(63a)和(63b)表明一般地对基于离散小波变换的谱, 这不是真的). 为证明相应的细节不能以类似

的方式使用，我们还计算

$$\widehat{\operatorname{var}}\{\widetilde{\mathcal{D}}_j\}\equiv\frac{1}{N}\|\widetilde{\mathcal{D}}_j\|^2=\frac{1}{N}\|\mathcal{T}^5\widetilde{\mathcal{D}}_j\|^2,\quad j=1,\cdots,7$$

且在第一行的后两个图中为 $\mathbf{X}$ 和 $\mathcal{T}^5\mathbf{X}$ 绘出了图. 练习[5.7]已经断言 $\|\widetilde{\mathcal{D}}_j\|^2\leqslant\|\widetilde{\mathbf{W}}_j\|^2$，因此我们一定有 $\widehat{\operatorname{var}}\{\widetilde{\mathcal{D}}_j\}\leqslant P_{\widetilde{W}}(\tau_j)$. 图181表明在这里这是正确的. 关于 $\mathbf{X}$ 的能量特性是如何在 τ_j 的一个尺度下改变的研究应该建立在 $\widetilde{\mathbf{W}}_j$ 上而不是建立在 $\widetilde{\mathcal{D}}_j$ 上.

虽然极大重叠离散小波变换细节不能用来产生一个有效的方差分析，但它们是加性多分辨分析分解的一个整体部分，对于脉冲序列和它的平移，我们有

$$\mathbf{X}=\sum_{j=1}^{4}\widetilde{\mathcal{D}}_j+\widetilde{\mathcal{S}}_4\quad 和\quad \mathcal{T}^5\mathbf{X}=\sum_{j=1}^{4}\mathcal{T}^5\widetilde{\mathcal{D}}_j+\mathcal{T}^5\widetilde{\mathcal{S}}_4,$$

因此，把在图181后两列的中间五行序列加起来产生底行的时间序列. 对这个例子，

$$\mathbf{X}\neq\sum_{j=1}^{4}\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j+\mathcal{T}^{-|\nu_4^{(G)}|}\widetilde{\mathbf{V}}_4,$$

所以，我们一般不能直接用极大重叠离散小波变换系数产生一个加性分解(然而，练习[5.10]表示哈尔离散小波变换是一个例外，它的系数能用来直接产生一个加性分解). 用来产生 $\widetilde{\mathcal{D}}_j$ 和 $\widetilde{\mathcal{S}}_{J_0}$ 的滤波器的精确零相位特性对调准原始序列里和多分辨分析里的那些结果是很重要的，并且由多分辨分析提供的精确加性确保了 $\mathbf{X}$ 里所有的特性用细节或光滑来解释.

对于方差分析，加性多分辨分析的使用和极大重叠离散小波变换系数的使用之间的相互作用在下面将进一步阐述，其中我们考虑实际时间序列的基于极大重叠离散小波变换的小波分析.

5.7 例子：心电图数据

这里我们用极大重叠离散小波变换再检验一下心电图时间序列，主要目的是将它与4.10节用离散小波变换的分析进行比较. 图183表示基于LA(8)小波滤波器的 $J_0=6$ 层的极大重叠离散小波变换——相应的离散小波变换分析在图127中表示. 绘图之前，系数向量 $\widetilde{\mathbf{W}}_j$ 和 $\widetilde{\mathbf{V}}_6$ 分别循环前移式(114b)中的 $|\nu_j^{(H)}|$ 和式(114a)中的 $|\nu_6^{(G)}|$，尽可能使系数与原时间序列调准(平移的数量由图中记录的每个 $\mathcal{T}$ 的幂表示). 两条粗垂线穿越给定的 $\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j$ 和 $\mathcal{T}^{-|\nu_6^{(G)}|}\widetilde{\mathbf{V}}_6$ 描绘了被循环性假设至少在一定程度上影响的平移向量的开始和末端的区域(我们将在5.11节讨论怎样计算出这些直线的位置). 极大重叠离散小波变换系数审慎的子抽样和再尺度化产生离散小波变换系数，如果我们将图127里比较大尺度的系数与图183中的系数进行比较，就能看到离散小波变换里固有的予抽样相当重要地限制了我们

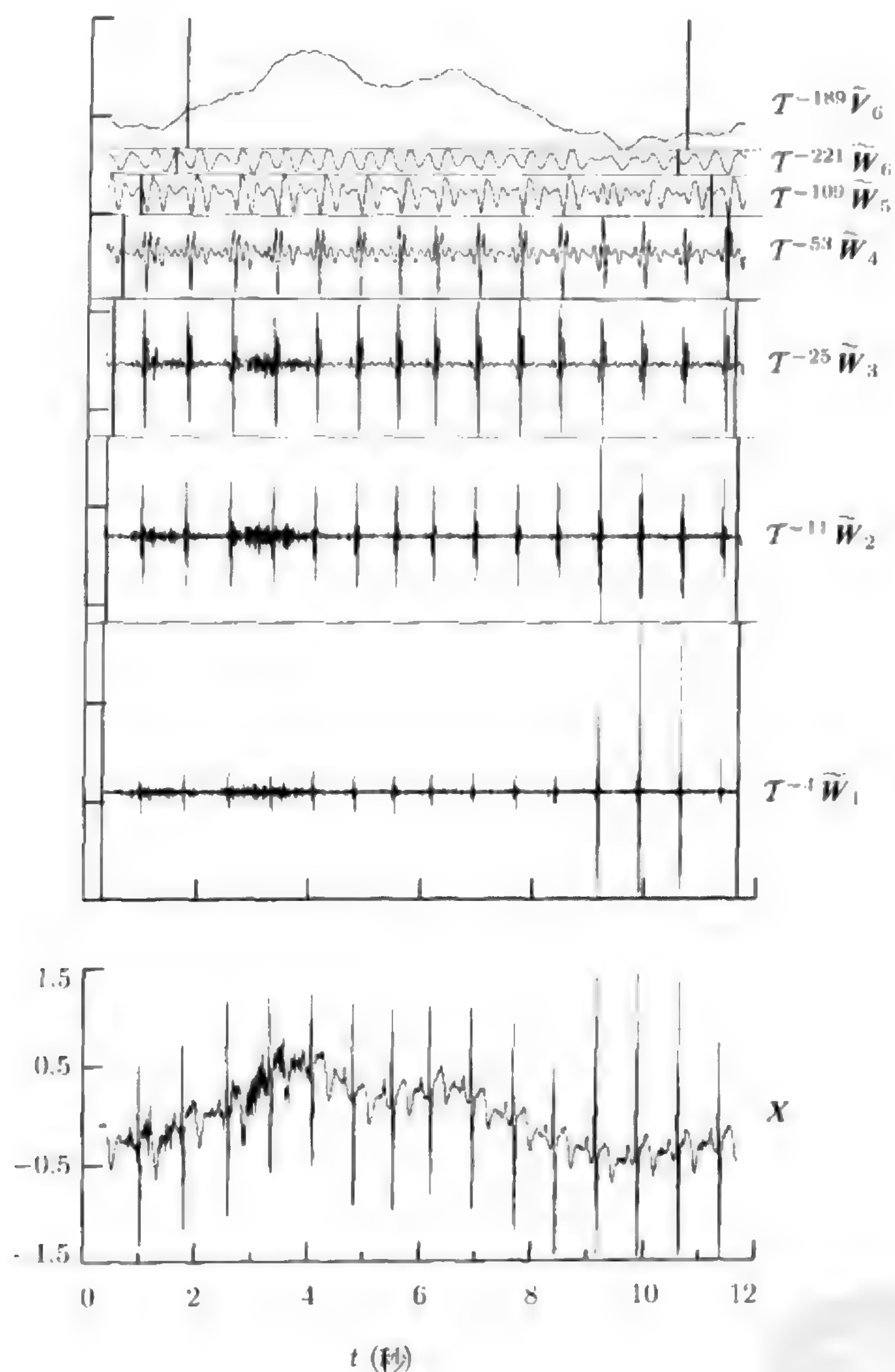


图 183　用 LA(8) 小波对心电图时间序列的极大重叠离散小波分析．上图应该和图 127 所对应的离散小波变化比较(详情参见正文)

按照原心电图序列里的结果解释其系数的能力．

在图 184 中表示出了极大重叠离散小波变换多分辨分析．和以前一样，粗垂线穿过给定的 $\widetilde{D}_j$ 或 $\widetilde{S}_6$ 描绘了被循环性假设至少在一定程度上影响的平移向量的开始和末端的区域(详情参见 5.11)．这个多分辨分析应该与图 133 中表示的基于

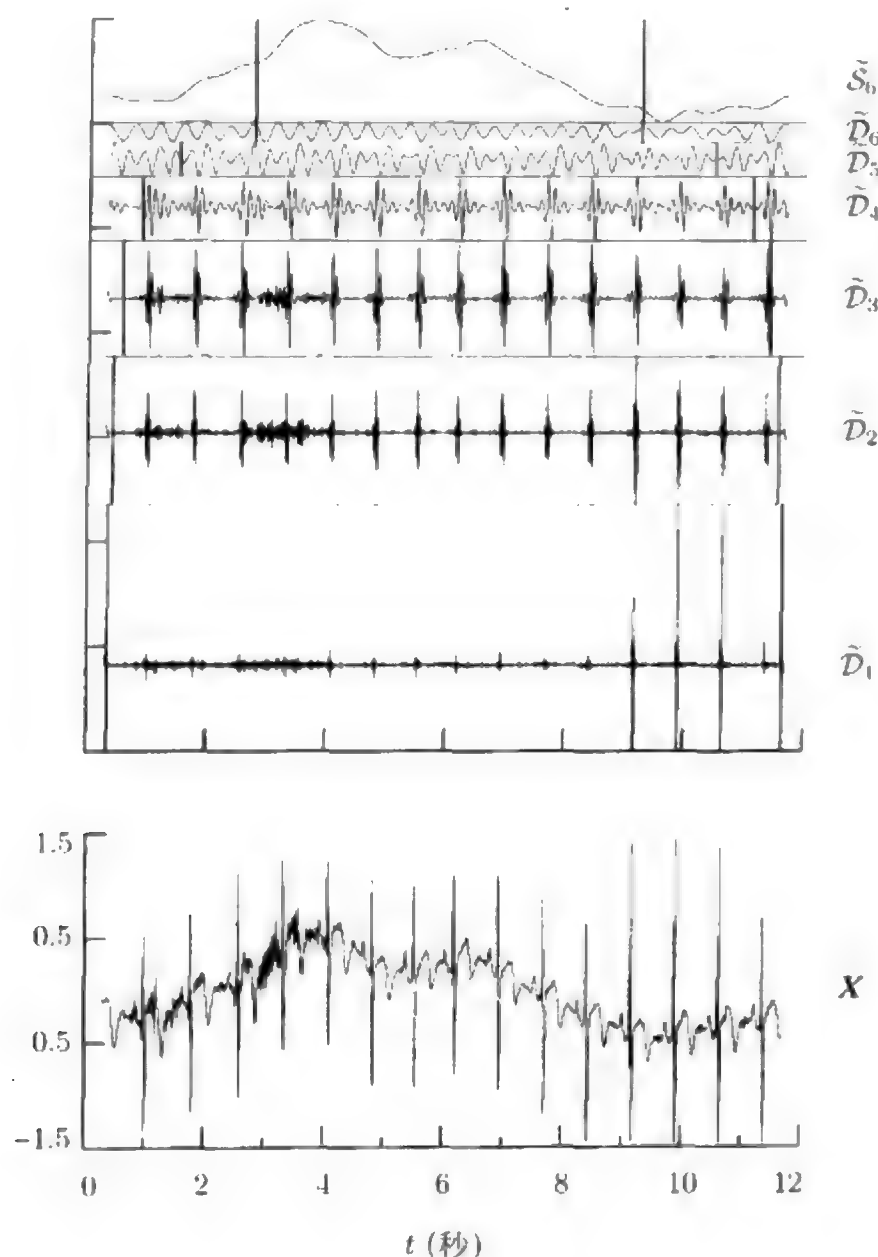


图 184 用 LA(8)小波对心电图时间序列的极大重叠离散小波变换多分辨分析(详情参见正文). 图 133 表示相应的离散小波变换分析

离散小波变换的多分辨分析进行比较. 这两个多分辨分析的总体概貌是很类似的，但注意在离散小波变换细节 $\mathcal{D}_3$、$\mathcal{D}_4$ 和 $\mathcal{D}_5$ 与相应的极大重叠离散小波变换细节 $\widetilde{\mathcal{D}}_3$、$\widetilde{\mathcal{D}}_4$ 和 $\widetilde{\mathcal{D}}_5$ 之间的差别：在每种情况中，极大重叠离散小波变换细节跨越时间时更一致，表明在离散小波变换细节上的时间变化在一定程度上归功于由离散小波变换假设的特定的校准. 还要注意极大重叠离散小波变换细节 $\widetilde{\mathcal{D}}_6$ 不和离散小波变换细节 $\mathcal{D}_6$ 一样逐渐衰减(尤其是在序列的开始)，毫无疑问，又是因为离散小波变换中的定位的影响. 在 $\mathcal{D}_6$ 里特征化“脉冲”的非对称状态也由于与此细节相关的零相位特性而在极大重叠离散小波变换细节 $\widetilde{\mathcal{D}}_6$ 里被消除了.

最后，让我们比较对 183 中的极大重叠离散小波变换小波和尺度系数与图 184 中极大重叠离散小波变换细节和光滑. 虽然前者保持了能量而后者保持了可加性，但这两个尺度分解的总体形式是十分相似的. 细节和光滑是通过滤波小波和尺度系数得到的，因此每个 $\widetilde{\mathcal{D}}_j$ 或 $\widetilde{\mathcal{S}}_6$ 在形式上必然比相应的 $\widetilde{\mathbf{W}}_j$ 或 $\widetilde{\mathbf{V}}_6$ 光滑. 这在 $\widetilde{\mathbf{V}}_6$ 与 $\widetilde{\mathcal{S}}_6$ 的比较中尤其明显：后者包含了前者里几乎完全缺少的小波纹. 虽然 $\widetilde{\mathcal{S}}_6$ 是基线偏移的一个相当有用的估计(同样地，在其他时间序列里，$\widetilde{\mathcal{S}}_{J_0}$ 是一个有用的跳动的估计)，任何在心电图序列里通过尺度和(或)时间的有意义的能量分划必须用极大重叠离散小波变换系数做. 因此两种类型的分析($\widetilde{\mathbf{W}}_j$ 与 $\widetilde{\mathbf{V}}_{J_0}$ 和 $\widetilde{\mathcal{D}}_j$ 与 $\widetilde{\mathcal{S}}_{J_0}$)在这个和其他时间序列的分析中有着重要的作用.

5.8 例子：子潮海平面涨落

在这一节中，我们用位于北加利福尼亚的开放海岸上伊斯兰城(Crescent City)的子潮海平面的时间序列来解释基于极大重叠离散小波变换的多分辨分析和方差分析. 在它的港湾内，一个永久的海浪测量仪器由国际海洋协会(NOS)保存. 这个测量器在一个平静的井中每隔 6min 测量一次水平面，这抑制了由于风波和上涨引起的高频涨落. 周期化水平调查保证了测量器的参考平面相对于周围的地方是保持恒定的. 这里考虑的时间序列是基于一段从 1980 年初到 1991 年末的水平面序列. 这个部分除了几个相对小又短的缺口之外还包括了长度为两周的一个缺口. 这两个周的缺口发生在 1990 年夏天且已经用预测潮加局部均值被填充，然而由国际海洋协会提供的插值用于接通短的缺口. 然后编辑的序列每 10 个值被下取样来产生每小时的值，且用 Kaiser 低通滤波器滤波来除去每天(一天)和半天(每天两次)的海浪. 滤波器平方增益函数的 3 dB 下(半幂)的点发生在对应于 1.82 天的周期的一个频率上，而对应于一个 2.68 天的周期和更长的频率是在低通滤波器的通带中，且在能量上以少于 1% 被截断. 最后，滤波的序列每隔 $\Delta t = 1/2$ 天重新采样，产生一个用厘米来测量的含有 $N = 8\,746$ 个值的时间序列 $\{X_t\}$. 这个序列反映了在伊斯兰城的沿海水平面上的子潮变化(即可预测潮成分消去后的变化)，且在图 186 的最下面绘出(关于这个应用的更详细的内容，参见 Percival and Mofjeld，1997).

描绘 $J_0 = 7$ 层 LA(8)基于极大重叠离散小波变换的多分辨分析的八个其他序列，在图 186 中伊斯兰城海平面数据的上面绘出. 最上面的序列是极大重叠离散小波变换光滑 $\widetilde{\mathcal{S}}_7$(相应于物理尺度 $\lambda_7 \Delta t = 64$ 天上的平均值)，而在它下面的 7 个序列是极大重叠离散小波变换的细节序列 $\widetilde{\mathcal{D}}_1$，$\widetilde{\mathcal{D}}_2$，…，$\widetilde{\mathcal{D}}_7$(相对应于从最下面到最上面依次为 1/2，1，…，32 天的物理尺度上的变化). 如往常，我们有 $\mathbf{X} = \sum_j \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_7$. 如 5.11 节所解释的，两条相交实垂线给出的 $\widetilde{\mathcal{D}}_j$ 或 $\widetilde{\mathcal{S}}_7$ 描绘了在某些程度上(尽管很小)用假设 $\{X_t\}$ 对 $t<0$ 和 $t \geq N$ 是循环定义的，这会影响按年的多分辨分析元素的区域；然而，这个序列的一个循环扩展在这里是有道理的，

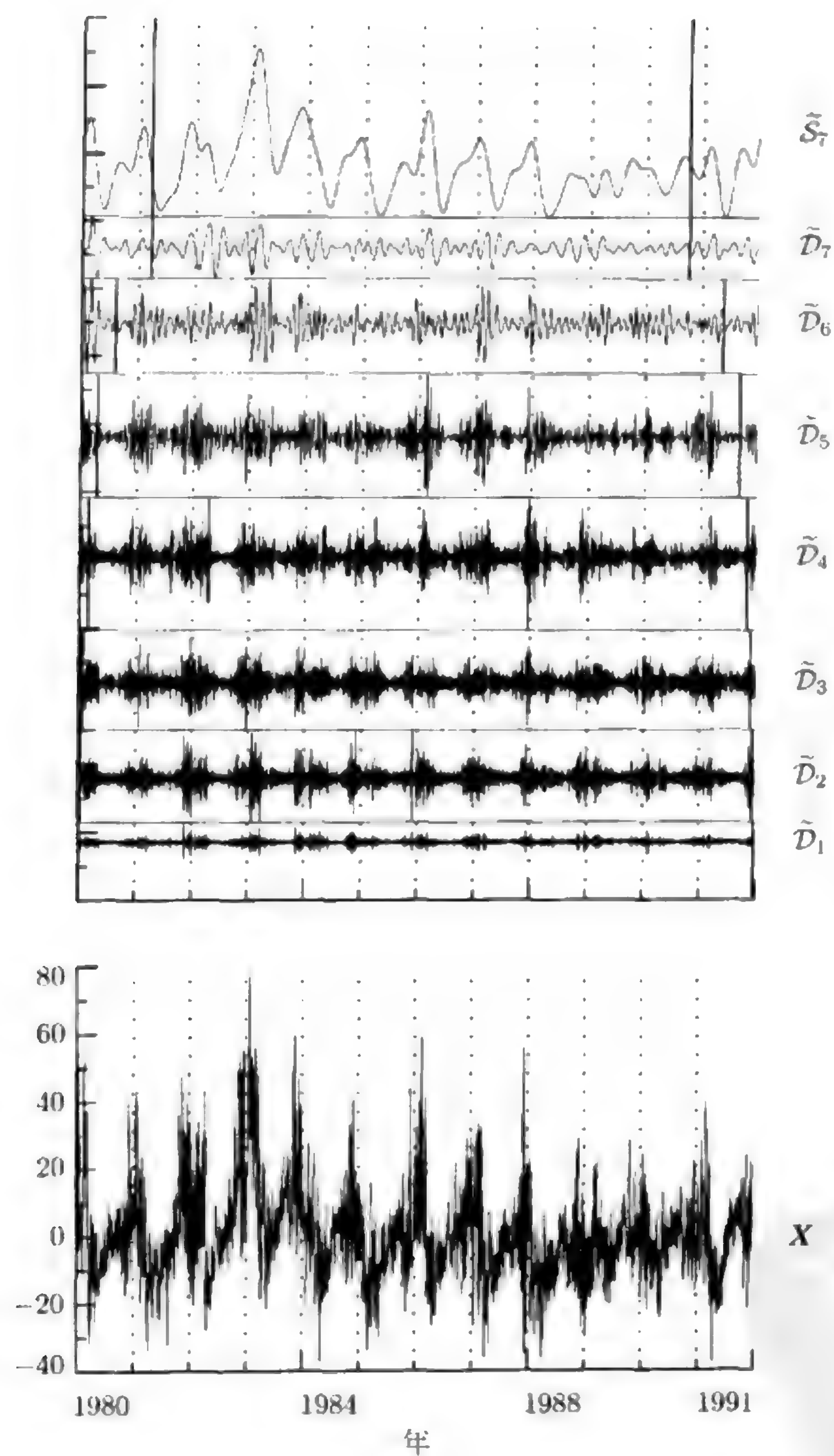


图 186 关于伊斯兰城子潮变化的 LA(8)极大重叠离散小波变换多分辨分析(详情参见正文)，这个序列是用厘米测量的

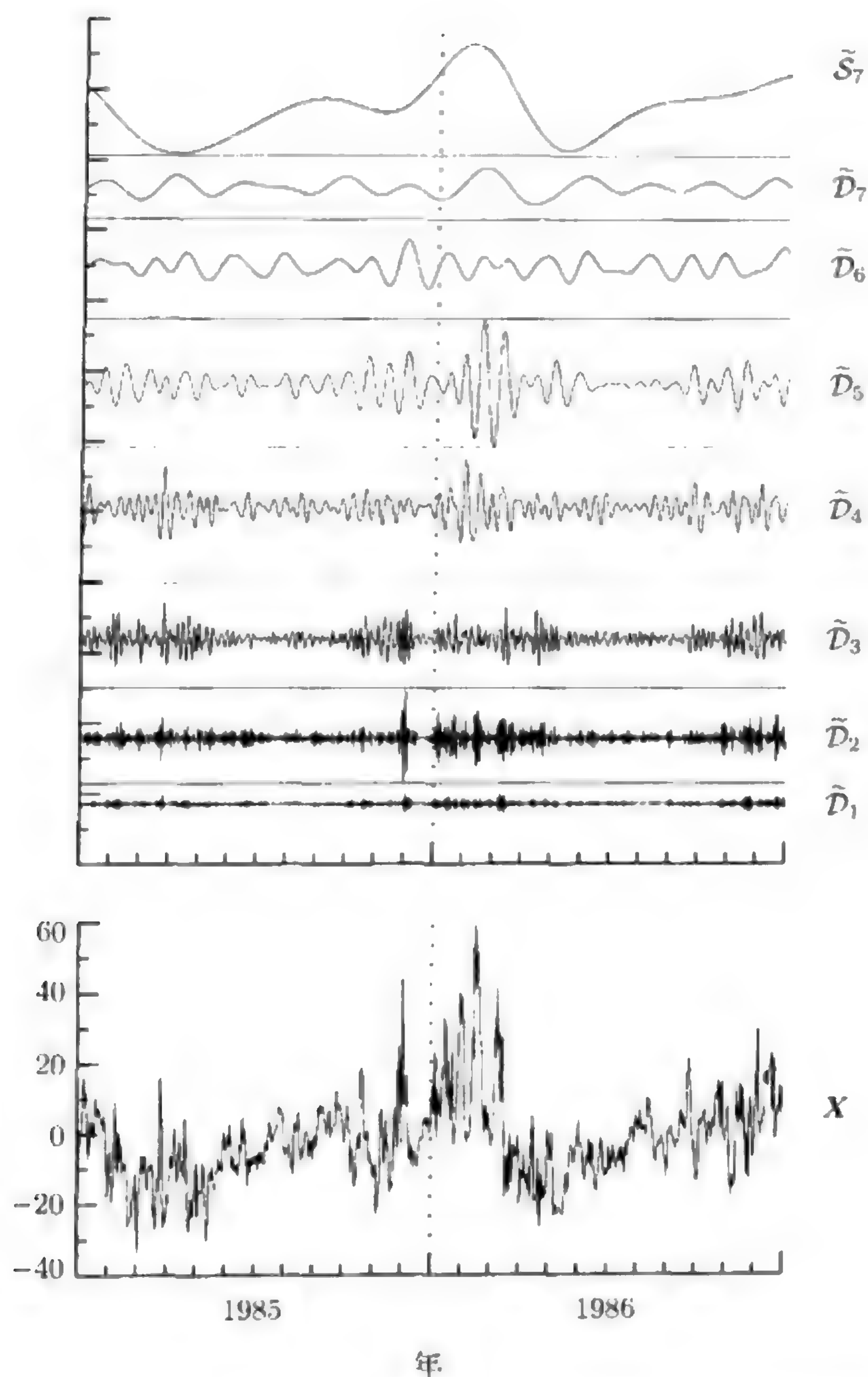


图 187 图 186 的 1985 年和 1986 年部分的扩大图

我们在观察现象时，有非常接近于 12 个完整年的数据量.

然而，在更详细地考虑多分辨分析之前，让我们表明为何决定用 LA(8) 小波滤波器，并且设 $J_0=7$. 选择这个特别的滤波器是因为我们最终想要把极大重叠离散小波变换系数本身与实际时间相结合（如果我们仅对多分辨分析感兴趣，D(4) 滤波器将会产生几乎相同的结果，但是对基于哈尔滤波器的小尺度，细节序列遇到一个“过滤”的形式. 因为显而易见地它们在跟踪更高尺度的波动——参见 8.9 节对于过滤的讨论和在各种小波滤波器中它是怎样被嵌入微分算子的数目

影响的). 最接近对称滤波器的相位形式允许我们与时间结合，因此，我们选取这种类型的滤波器且特别地停留在LA(8)滤波器上，这是因为$L=10$和12的最接近对称滤波器几乎产生了相同的分析，但是必须更多地考虑到由边界条件影响的项的数目(参见5.11节). coiflet滤波器也是合理的选择，但是这种类型的最短成员——C(6) coiflet——给出了一个相似的多分辨分析，当最短尺度上的小波系数用于执行将在8.7节中讨论的子序列方差分析时，也有它泄漏的证据(下一个最短coiflet是C(12)，它给出了与LA(8)滤波器相同的结果，但是与LA(10)和LA(12)滤波器有同样的关于边界的影响).

我们选择$J_0=7$层是建立在如下事实的基础上的：得到的光滑$\widetilde{S}_7$反映了在一个物理尺度$\lambda_7\Delta t=2^7\dfrac{1}{2}=64$天上的平均值，因此包含了在子潮序列中的年内的变量(即年之间的变量). 如果我们使$J_0>7$，就会将这些年变量打破成为碎片，这些与分离的统一体相比对海洋学者而言是不感兴趣的. 另一方面，如果设为$J_0<7$，就会将某些尺度加到有一个季节性的依赖变量$\widetilde{S}_{J_0}$上(下面我们要看的)，因此可以更好地分开研究.

现在让我们更详细地考虑多分辨分析. 对于多分辨分析的一个用途是检测波动，即本质上是瞬变的而且限制到几个尺度上的波动. 作为一个例子，图187给出了图186一部分的一个扩展观察，并且表明在变化上的一个短的间歇发生在1985年12月(恰恰在1986年初标记过的垂直虚线的左边). 这次间歇发生在最短尺度上，即1、2和4天(相应于$\widetilde{\mathcal{D}}_2$、$\widetilde{\mathcal{D}}_3$和$\widetilde{\mathcal{D}}_4$). 这期间的气候条件的检查表明间歇是由于风暴的向北偏移，这是在这段时间中由在美国西部形成的空气中一个高压的脊向远离伊斯兰城区形成的. 如在较长尺度中可看到的一样(8、16和32天，相对应于$\widetilde{\mathcal{D}}_5$、$\widetilde{\mathcal{D}}_6$和$\widetilde{\mathcal{D}}_7$)，这样的脊慢慢改变了空气压力模式和风模式. 然而，这个脊不能存留足够长来反转在光滑$\widetilde{S}_7$上可见的季节性水平面上的向上倾向(图187最上面的序列)，即使它的两个高空压力和在海岸附近北美风易于产生低水平面. 它对伊斯兰城水平面的主要影响以在较短尺度上变化性的递减出现. 回顾过去，这种递减能在图187最下面的$\{X_t\}$的图中看到，但是在图中认出来是很困难的，然而多分辨分析使它很容易挑选出海洋学者感兴趣的这样和那样的特点.

在图186中多分辨分析也使得一些尺度上的变化很明显是季节性相关的. 然而，从这个分析中，定量研究$\{X_t\}$的尺度相关变化特性是有问题的，因为这里

$$\|\mathbf{X}\|^2\neq\sum_{j=1}^{7}\|\widetilde{\mathcal{D}}_j\|^2+\|\widetilde{S}_7\|^2.$$

我们可以用

$$\|\mathbf{X}\|^2=\sum_{j=1}^{7}\|\widetilde{\mathbf{W}}_j\|^2+\|\widetilde{\mathbf{V}}_7\|^2$$

替代检查极大重叠离散小波变换系数. 我们可以通过如下定义的创造旋转积累方差图，定量研究在一个尺度对尺度基础上的方差的时间相关性. 首先我们循环地推进$\widetilde{\mathbf{W}}_j$ 的元素 $|\nu_j^{(H)}|$ 单位，以使 $\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j$ 与 $\mathbf{X}$ 一致(参见式(114b)和围绕它的讨论). 然后我们渐增地求 $\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j$ 的元素的平方和用 $1/N$ 规范化以形成

$$C_{j,t} \equiv \frac{1}{N}\sum_{u=0}^{t}\widetilde{W}^2_{j,u+|\nu_j^{(H)}| \bmod N}, t = 0,\cdots,N-1.$$

由于 $C_{j,N-1}=\frac{1}{N}\|\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j\|^2=\frac{1}{N}\|\widetilde{\mathbf{W}}_j\|^2$ 能够认作是 $\mathbf{X}$ 的由物理尺度 $\tau_j\Delta t$ 的抽样方差的分布，序列 $\{C_{j,t}\}$ 对通过时间的第 j 个尺度，跟踪抽样方差的积聚. 因为 $\{C_{j,t}\}$ 是一个非减的序列，有时候作为对 t 的函数绘出 $C_{j,t}$ 来积聚抽样方差的这种方式是很困难的，但是我们能够通过绘出

$$C'_{j,t} \equiv C_{j,t} - t\frac{C_{j,N-1}}{N-1}, \quad 作为 t 的函数 \tag{189}$$

克服这个困难，我们称之为旋转积累方差图(本质上，旋转允许我们看到，按t 的每个增量$\frac{C_{j,N-1}}{N-1}$的一致积聚形成的偏差；这种观点将在 9.6 节被重新探索以估计在一个时间序列变量上的变化的位置).

作为例子，图 190 展示了对物理尺度 $\tau_2\Delta t=1$ 天和 $\tau_7\Delta t=32$ 天上旋转积累方差图(因为对于 LA(8)的小波滤波器，$\nu=-3$，式(114b)告诉我们 $\nu_2^{(H)}=-11$ 和 $\nu_7^{(H)}=-445$). 在这个图中最下面的图很明显地表明，在天的尺度上变化的可变性是季节性相关的，在冬季发生更多的变化(即在垂直虚线的附近). 在夏季月份里，旋转积累方差图的线性表示表明，在仲夏的变化性是相当稳定的，而且从一年到另一年是可以预测的. 这表明稳定的夏季模式改变到更飘忽不定的冬季模式，可以从一年变化到下一年，达到一个月之多；进而，这改变看起来好像是某年(1985)清楚地突变而在其他年(1998)分散. 另一方面，图 190 最上面的图表明在 32 天尺度上变化的可变性的相应图形. 在这个尺度上的季节依赖性是更不确定的. 例如，旋转积累方差表明，在 1980、1982、1983、1986、1987 和 1991 初附近上变化的增长被标记，但是从大约 1989 年 3 月到 1991 年 1 月，大约有 20 个月的跨度几乎没有变(在 8.7 节我们将进一步考虑这个旋转积累方差的时间变化方面).

最后，如我们在 5.3 节所注意的，由极大重叠离散小波变换构造的细节和光滑能够认作是零相位滤波器的输出. 如果想要把细节里的结果和原始序列里的结果联系起来，而且要理解不同尺度的结果之间的暂时关系，那么这个零相位特性是很重要的. 为了阐述这个零相位特性在实际中意味着什么，我们在图 191 最上面的图中表示出了极大重叠离散小波变换细节 $\widetilde{\mathcal{D}}_5$(细曲线)和 $\mathbf{X}$(粗曲线)在 1986

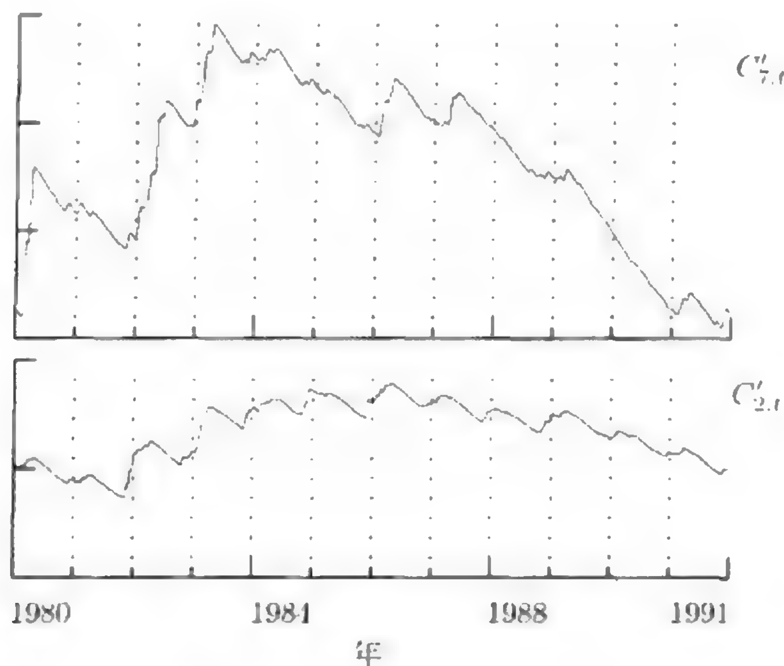


图 190　子潮海平面变化在物理尺度 $\tau_2\Delta t=1$ 天(底图)和 $\tau_7\Delta t=32$ 天(顶图)的旋转累积方差图

年的第一部分附近. 这个细节应反映 $\boldsymbol{X}$ 在尺度 8 天上的变化，因此我们期待当在大约 8 天上的一个平均值不同于在它附近标记过的值时，在 $\widetilde{\mathcal{D}}_5$ 上的值是大的. 选择这个时间周期是因为在 $\boldsymbol{X}$ 中有这样持续时间的三个尖峰——它们的中心在从 1986 年初大约 32、48 和 72 天(注意在横轴上，小的标记记号之间的距离是 8 天). 极大重叠离散小波变换细节很好地挑选出这些结果，而且在 $\widetilde{\mathcal{D}}_5$ 中尖峰的位置严格地与 $\boldsymbol{X}$ 中的那些相匹配. 因为当我们对 $\boldsymbol{X}$ 运用循环平移时，$\widetilde{\mathcal{D}}_5$ 也循环平移，无论我们怎样地选择 $\boldsymbol{X}$ 的开端，最上面的图形将保持相同.

为了比较，图 191 的中间和下面的图形中的细曲线表明，对于在 $\boldsymbol{X}$ 开端两个不同平移的离散小波变换细节 $\widetilde{\mathcal{D}}_5$，即平移 12 和 8 值(相应于 6 天和 4 天). 注意用两天改变 $\boldsymbol{X}$ 的开端将显然地改变 $\widetilde{\mathcal{D}}_5$. 因为离散小波变换细节不保持零相位特性，$\widetilde{\mathcal{D}}_5$ 中峰值的位置就不必与 $\boldsymbol{X}$ 中的那些匹配很好. 例如，考虑 $\widetilde{\mathcal{D}}_5$ 中尖峰的位置相对应于 $\boldsymbol{X}$ 中 32 天附近. 与上面图形里的 $\widetilde{\mathcal{D}}_5$ 相关，在 $\widetilde{\mathcal{D}}_5$ 中的尖峰明显发生在中间图形的后端和在最下面图形的前端. 离散小波变换细节定位于 $\boldsymbol{X}$ 中瞬时结果的能力就这样被限制，然而因为它们的零相位性质，极大重叠离散小波变换细节表现得很好.

5.9　例子：尼罗河最低水位

这里我们提出一个时间序列的多分辨分析，这个序列是由开罗(Cairo)附近的 Roda 测定的数据，记录了从 622 年到 1284 年的尼罗河年最低水位的测量组成(这个分析是从华盛顿大学统计系的 Simon Byers 未出版的 1995 年的报告中摘取的). 以米测量的水位的时间序列在图 192 的最下面绘出，且实际上是尼罗河最

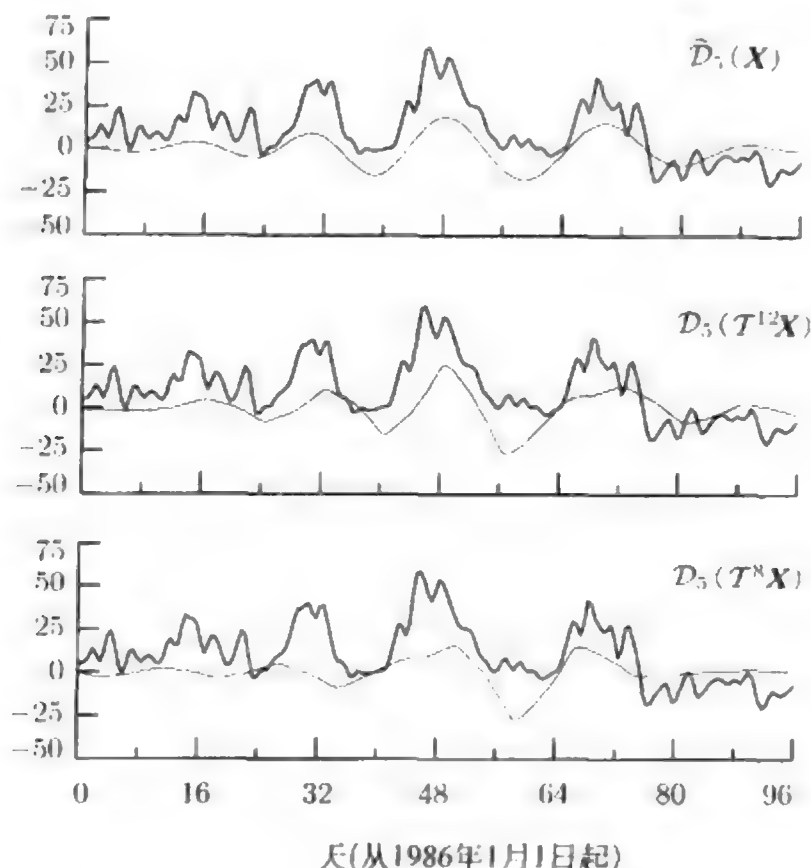


图 191　对极大重叠离散小波变换细节的零相位特性的说明．每个图的粗曲线是在 1986 年初开始的伊斯兰城序列 96 天的那部分．细曲线(从顶部到底部)是原时间序列 $\boldsymbol{X}$ 的 LA(8)极大重叠离散小波变换细节 $\widetilde{\mathcal{D}}_5$，$\mathcal{T}^{12}\boldsymbol{X}$(延迟 12 个数据值的时间序列)的 LA(8)离散小波变换细节 $\mathcal{D}_5$ 和 $\mathcal{T}^{8}\boldsymbol{X}$(延长 8 个数据值的时间序列)的 LA(8)离散小波变换细节 $\mathcal{D}_5$．与这些细节相关的物理尺度是 $\tau_5\Delta t=8$ 天；即细节应该与一个 8 天尺度上平均值的变化有关(这就是横轴上较小点标号间的距离)

低水位的一个非常长的历史记录的一部分，而这是由 Prince Omar Toussoun 和 1925 年出版的多卷《Mémoires de l'Institut d'Égypte》中标题为《Mémoire sur L'Histoire du Nil》中的资源进行编辑的．Toussoun 的记录扩展到 1921，从 622 年到 1284 年表示最长一段无间断的值(1284 年以后有一些小的分散的丢失值直到 1470 年，之后覆盖就更不完全了：例如在整个 16 世纪只有 17 个测量值)．阿斯旺・坝(1902)和阿斯旺・高坝(1960～1970)的建造结束了自然界中控制尼罗河水流的局面．

虽然对于在图 192 中的时间序列负责的主要现象不再是可操作的，尼罗河最低水位的历史记录以它特有的正确性发展了一段很有兴趣的历史．水文工作者 Hurst(1951)用这个和其他的时间序列在一次研究中演示出传统的统计方法对时间序列的无效性，且展示了长项的持续性(即慢衰减的自相关序列——参见 7.4 节)．在 Hurst 的先驱性工作之后，Mandelbrot and Wallis(1968)提出用分形高斯噪声(这将在 7.6 节中定义)来模型化他所研究的各种序列以描述 Hurst 的经验

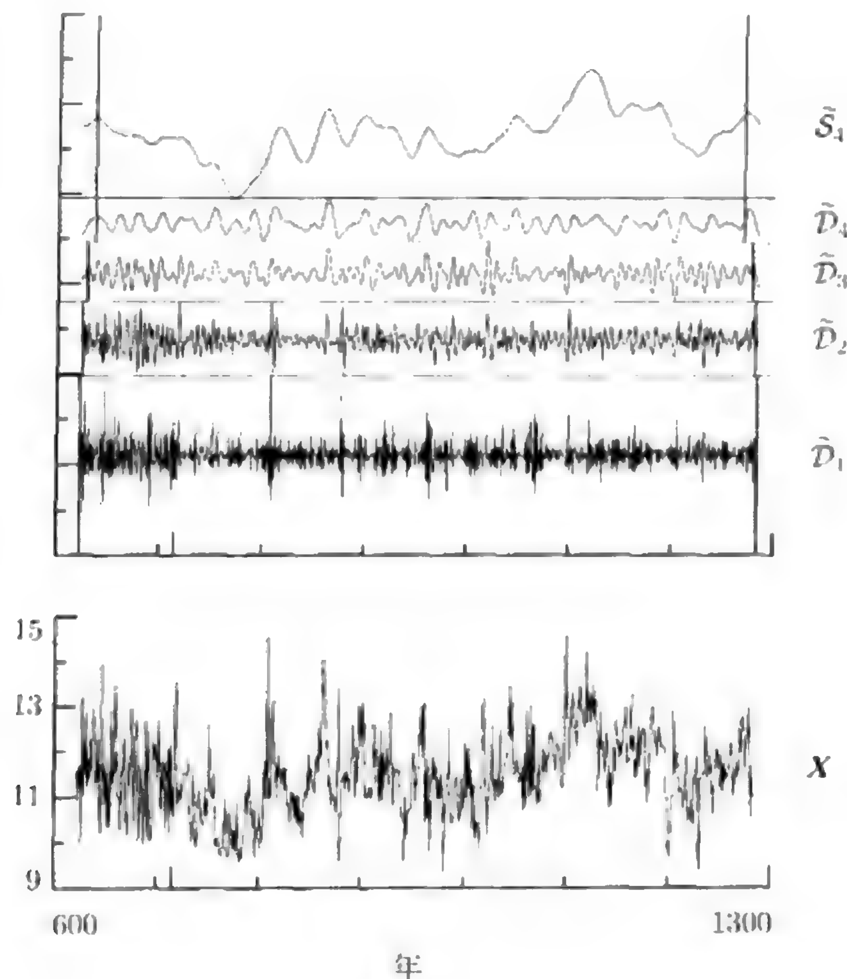


图 192 利用哈尔小波滤波器关于尼罗河最低水位(用米)的 $J_0=4$ 层极大重叠离散小波变换多分辨分析

结果. 这个模型首先完全地试图在特征化现在已经很著名的稳定长记忆过程(第 9 章观察了小波可用于研究这种过程的各种方法). Beran(1994)在他的开拓性的有关长记忆过程的统计学书中，使用扩展的尼罗河序列作为一个例子. 最近的 Eltahir and Wang(1999)以确定历史上厄尔尼诺(El Niño)现象多久发生的间接方法来研究这个序列.

图 192 的上面部分展示了对尼罗河的 $J_0=4$ 层的极大重叠离散小波变换多分辨分析. 我们使用哈尔小波滤波器，因为从 $L=4$ 或更宽的滤波器中的多分辨分析与哈尔多分辨分析在数量上类似——既然用更宽的滤波器没有明显可见的好处，那么哈尔滤波器在这里是更可取的，因为它最小化了有关边界条件效果的顾虑(我们用一对粗垂线来表示每个多分辨分析分量的边界区域——参见 5.11 节). 选择 $J_0=4$ 是因为我们主要在指出通过尺度 τ_1 和 τ_2 的确定时间依赖性，把它和在更高尺度上缺乏这样依赖性相比较感兴趣. 由于这个数据是按年记录的(即 $\Delta t=1$ 年)，多分辨分析的分量连带从最下面到最上面依次为 1、2、4 和 8 年的物理尺度上的平均的变化(分别从 $\tilde{D}_1$ 升到 $\tilde{D}_4$)和 16 年的平均($\tilde{S}_4$). 多分辨分析最感兴

趣的方面是在细节$\widetilde{\mathcal{D}}_1$和$\widetilde{\mathcal{D}}_2$中变化的明显不均等：表面上看起来在这两个细节中直到早期的700年有比较大的变化性．对于较高尺度的细节($\widetilde{\mathcal{D}}_3$和$\widetilde{\mathcal{D}}_4$)，证明在记录的早期部分没有增加的变化性依据．当增加的变化性是很明显的——在一个较小的程度上——在$\{X_t\}$本身的图上且已经用其他方法研究过(Beran，1994)，基于小波的多分辨分析十分清楚地将它选出来，并且很好地证明了它是在尺度上被局部化的．Toussoun的字母组合及Popper(1951)和Balek(1977)后来的历史性研究都指出了715年(或其后不久)在开罗附近的尼罗河上Roda岛屿上的清真寺里“水位计”的构造(这年在图192的两个图上用横轴上的一个延长记号所标记)．从那个时候到1284年中每年的最低水位就用这个装置所测量(或是在861年做的一个新构造)．从622～714年测量的精确渊源是不被人所知的，但是他们最大可能在开罗附近的不同位置上，用Roda岛上清真寺的那个较差的不同类型的测量设备．因此这个新的水位计导致了最小尺度上变化性的减少，这种说法是很合理的(根据地球物理解释这种变化性的差别是很难的)．

我们将在8.8节和8.9节回到这个例子当中，在那里我们利用非正式的和正式的统计检验来核实在大约715年前后在一两天的尺度上的变化性是相当不同的．

5.10 例子：海洋切变测量

这里我们考查垂直海洋切变测量的一个“时间”序列．这些数据是通过越过船的侧面然后设计着垂直下降到大海里的方法而收集到的．随着下降，调查搜集了作为深度的函数的有关海洋的测量．因此，“时间”序列的次序变量就是深度．测量之一就是水速的x分量．这个速度的收集是每隔0.1米，先在10米的一个区间上分划，然后被低通滤波从而得到海洋里与垂直切变相关的一个序列．图194的底部图片展示垂直切变测量的序列在这里用到了(这个序列有秒分之一的单位)．数据从350.0米深延伸到1 037.4米，增量$\Delta t=0.1$m(总共有$N=6\ 875$个数据值)．在图上有两条标记过的细垂线，其间有从489.5米到899.0米范围变化的4 096个值(这一子序列将在8.9节中进一步讨论)．

图194表示出了用LA(8)小波滤波器构造的一个$J_0=6$层的多分辨分析．我们选择这个滤波器主要是因为，当我们在8.9节回到这个例子时，将会既对研究不同尺度的小波系数如何与时间有关感兴趣，又对能提供防备漏泄的小波滤波器感兴趣(漏泄的定义也将在8.9节讨论)．在时间上完成调整的自然选择是最接近对称(LA)或coiflet滤波器．在这两类滤波器中，LA(8)是具有可接受的漏泄特性的最短滤波器(如果我们只对多分辨分析感兴趣，那么D(4)、D(6)和C(6)滤波器将会产生与LA(8)滤波器的多分辨分析十分类似的多分辨分析，但是对于哈尔多分辨分析，小尺度细节$\widetilde{\mathcal{D}}_j$看起来由于漏泄而相当不同)．我们假设$J_0=6$是因为这个是明显不受破裂影响的最小尺度的指标(下面将会讨论)．这一选择将会产生使破裂隔离细节序列的一个多分辨分析，并且也会产生(为了比较)不受破裂

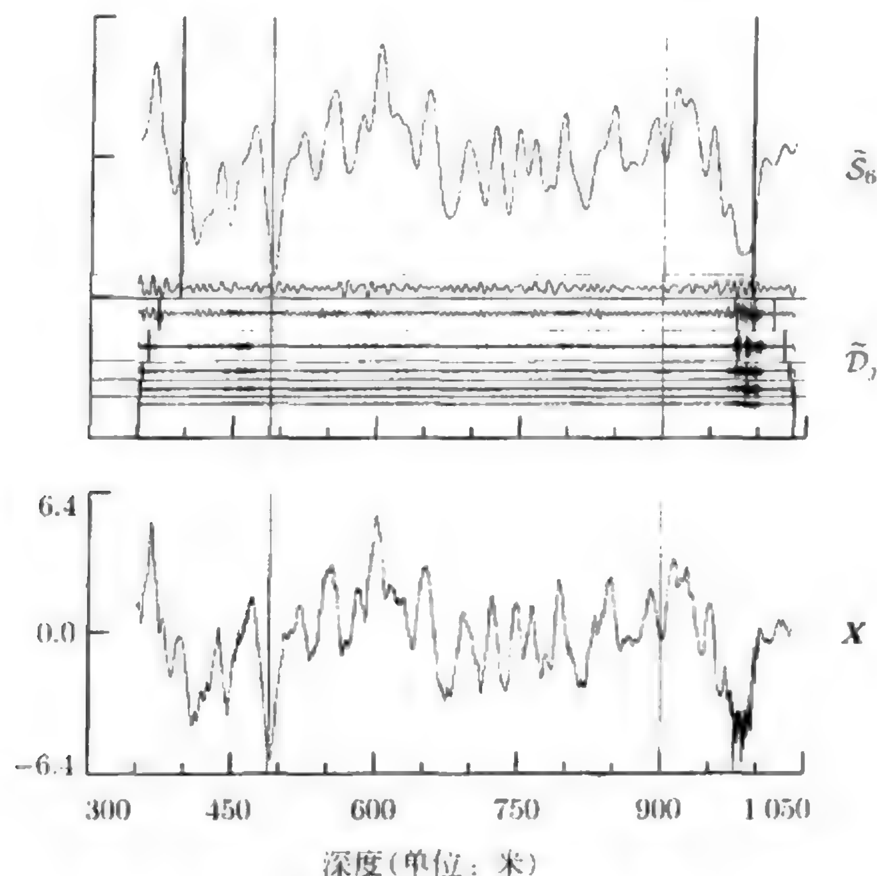


图 194 用垂直切变测量(用逆瞬间)的 LA(8)小波对深度(单位：米)的 $J_0=6$ 层极大重叠离散小波变换多分辨分析. 这个序列是由华盛顿大学应用物理实验室 Mike Gregg 搜集和提供的

影响的一个细节序列.

图 194 中的多分辨分析是用平滑 $\widetilde{S}_6$ 来支配的. 这个平滑在序列 $\{X_t\}$ 的除了开头和末尾之外的其他部分里，看起来是大尺度变化的一个合理的摘录，这里极大重叠离散小波变换里固有的周期假设可论证的原因曲解了在 $\widetilde{S}_6$ 上的失真(和以前一样，与给定的 $\widetilde{D}_j$ 或 $\widetilde{S}_6$ 相交的粗垂线对，绘出了至少在一定程度上受边界条件影响的多分辨分析的一部分). 如同需求的，我们能用映射边界条件来消除这些轻微的失真(参见围绕式(140)的讨论). 因为平滑支配着多分辨分析，且细节不是可辨别的，我们已经在图 195 里分别绘出了 $\widetilde{D}_j$. 细节很好地挑选出了两处破裂，一次在 450 米附近，另一次刚好在 950 米后(第二次破裂在图 194 里 $\{X_t\}$ 的图中可清楚地看到，但第一次仅在事后才看得明显). 因为与 $\widetilde{D}_j$ 相关的物理尺度是 $\tau_j\Delta t=0.1\tau_j$ 米，我们看到第一次破裂在 0.8 米和更小的尺度上是突出的(即 $\widetilde{D}_1$ 向上到 $\widetilde{D}_4$)，而在尺度 1.6 米(即 $\widetilde{D}_5$)上仍有第二次破裂的突出迹象，对 $J_0\geqslant 7$ 的多分辨的核查表明，对于长于 3.2 米的实际尺度没有这些破裂的迹象. 如果我们更仔细地检查第一次破裂，就会看到在 $\widetilde{D}_4$(0.8 米)处迅速下降，但破裂在越过这个下降点的较小尺度细节上($\widetilde{D}_3$、$\widetilde{D}_2$ 和 $\widetilde{D}_1$)延续. 一个物理解释可能是在一个 0.8 米尺度上的骚动驱使着在更深处较短尺度上的骚动. 相比较，第二大的破裂

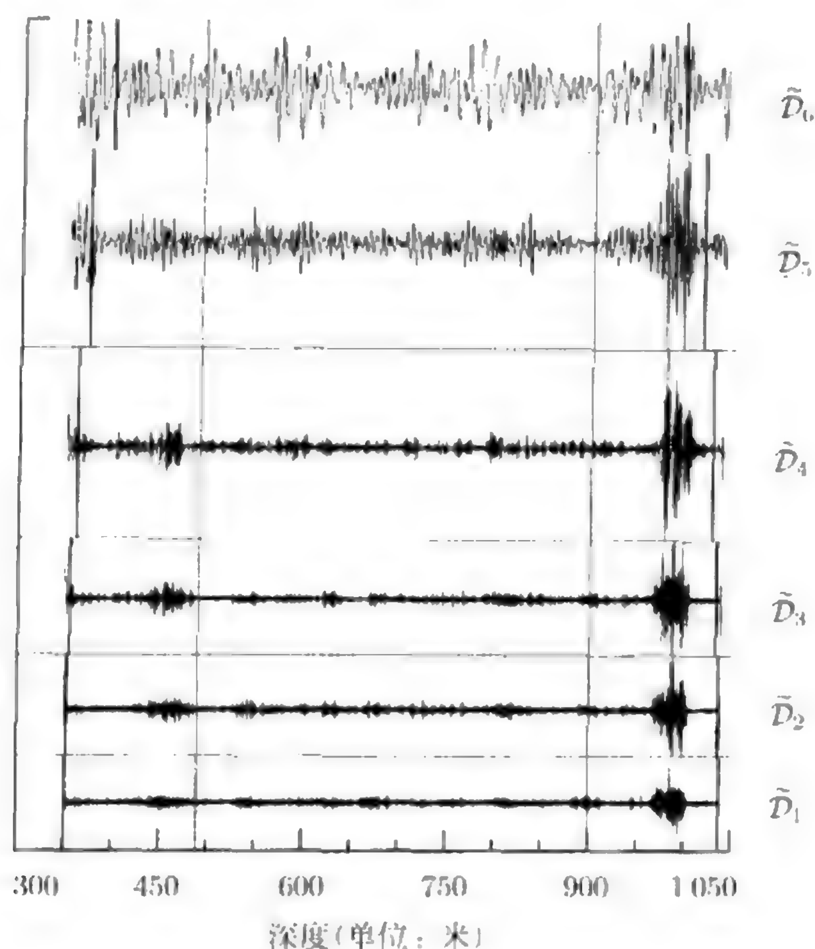


图 195　对图 194 表示的多分辨分析的细节序列的扩大图

看起来在细节 $\widetilde{D}_1$ 至到 $\widetilde{D}_4$ 被联合定位，但好像在 $\widetilde{D}_5$ 和 $\widetilde{D}_6$ 里，也有破裂的主要部分之前的增加变化性的线索. 一个极大重叠离散小波变换多分辨分析使得在不同尺度上的结果之间的深度(或时间)关系式是很重要的，因为它的细节和平滑与零相位滤波器相关.（虽然由多分辨分析给定的加性分解给出了切变序列的一个有用的基本描绘，但我们在已经讨论过用简单的统计模型评论基于小波的方差分析的方法后，将在 8.9 节里更多地说明关于它的特性.）

5.11　实际应用中需要考虑的问题

在 4.11 节中，我们讨论了实现时间序列的离散小波变换分析时必须考虑的几个实际因素，做的一些评论对极大重叠离散小波变换也是适合的；然而，由于离散小波变换和极大重叠离散小波变换之间有相当大的差别，故下列主题的每一个都值得更多的讨论.

• 小波滤波器的选择

在很大程度上，4.11 节做出的关于给离散小波变换选取一个特殊的小波滤波器的评论对极大重叠离散小波变换也是有效的，但如图 196 中所表述的，极大

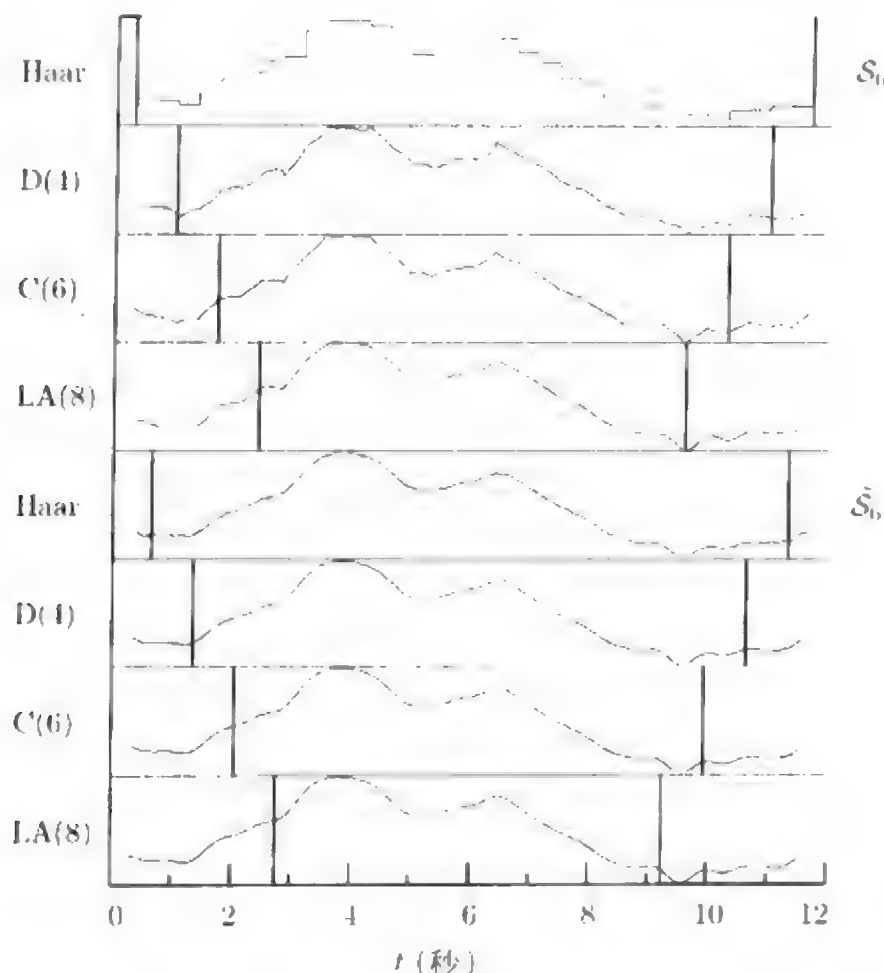


图 196 在每一组里从顶部到底部用哈尔，D(4)，C(6)和 LA(8)小波滤波器．对心电图时间序列离散小波变换平滑 S_6（顶部 4 个序列）和极大重叠离散小波变换平滑 $\widetilde{S}_6$（底部 4 个序列）的比较．粗垂线表明了平滑的边界区域，即，（至少在一定程度上）被循环性假设所影响的平滑

重叠离散小波变换对产生一个多分辨分析的用途使得对小波滤波器的选择更不重要了．图 196 顶部的 4 个序列表示了对心电图时间序列由哈尔、D(4)、C(6)和 LA(8)多分辨分析的离散小波变换为基础的平滑 S_6（这些取自于图 130 到图 133）．基于哈尔、D(4)和 C(6)小波滤波器的平滑包含了不引人注意的块、鱼翅和三角，这些是心电图数据里由于滤波器而非相关特性的典型产物．另一方面，LA(8)光滑表明没有明显的典型产物，且因此可以被认为是这个序列里基线偏移的更实际的描述．这些因素就建议我们对大多数时间序列不用哈尔、D(4)或 C(6)的基于离散小波变换的多分辨分析，因为没有很多序列，它的典型特性与块、鱼翅和三角相匹配．

现在，让我们看一下，当从离散小波变换转到极大重叠离散小波变换时将会发生什么．图 196 底部的 4 个序列表示出了建立在和以前相同的 4 个小波滤波器基础上的极大重叠离散小波变换光滑 $\widetilde{S}_6$．从性质上看，这三个光滑部分很相似，虽然哈尔 $\widetilde{S}_6$ 表示了其余 3 个极大重叠离散小波变换平滑所不具有的十分小的波纹的迹象．由于不同的小波滤波器，对心电图序列的极大重叠离散小波变换细节的一个

相似检验又一次表示出了相当小的差别. 因此，我们不能像离散小波变换实例中那样简单地处理哈尔、D(4)和C(6)多分辨分析. 一般地，为做出小波滤波器的一个合理的选择，我们必须借助于应用的相关考虑(例如，如果$\widetilde{S}_6$想获取基线偏差且如果为了论证，我们假设这个偏差的概念消除了小波纹，那么心电图分析家相对于哈尔将会更偏爱于D(4)、C(6)或LA(8)小波滤波器).

为什么基于极大重叠离散小波变换的多分辨分析比基于离散小波变换的多分辨分析更不依赖于小波滤波器的关键在练习[5.1]中给出，它将证明极大重叠离散小波变换细节和光滑，能通过平均某些从时间序列的循环平移形式中产生的循环平移离散小波变换细节和光滑而作用产生. 这个特别层次的平均使得块、鱼翅和三角很模糊，于是消除了基于离散小波变换的多分辨分析中的明显的缺陷. 从滤波理论的观点来看，产生了极大重叠离散小波变换细节$\widetilde{\mathcal{D}}_j$的有利特性，因为它是两级滤波器级联的输出(参见5.4节). 第一级涉及了具有传递函数$\widetilde{H}_j(\cdot)$的一个"分析"滤波器，而第二级是一个"综合"滤波器$\widetilde{H}_j^*(\cdot)$. 无论我们选择用什么样的小波滤波器，等价滤波器$|\widetilde{H}_j(\cdot)|^2$有一个总为零的相位函数. 这个解释暗示我们希望从不同小波滤波器的极大重叠离散小波变换小波系数$\widetilde{\mathbf{W}}_j$之间，找到比相应的细节之间更多的不同：$\widetilde{\mathbf{W}}_j$是用$\widetilde{H}_j(\cdot)$形成的，其相位特性依赖于特定的小波滤波器. 另外，与极大重叠离散小波变换小波系数和细节相联系的平方增益函数分别是$|\widetilde{H}_j(\cdot)|^2$和$|\widetilde{H}_j(\cdot)|^4$，因此我们希望后者能在某些前者可能不胜任的实例中，产生带通滤波器的一个还不错的逼近. 这点将由8.9节的例子阐述，其中证明了哈尔$\widetilde{\mathbf{W}}_j$的运用导致了一个与建立在长度$L \geq 4$的小波滤波器基础上的$\widetilde{\mathbf{W}}_j$相比，受到"泄漏"妨碍的方差分析.

作为总结，与基于离散小波变换的多分辨分析相比较，一个基于极大重叠离散小波变换的多分辨分析，更少依赖于我们对小波滤波器的选取，但不如说，我们总是推荐使用一个特定的滤波器. 建立在不同小波滤波器基础上的多分辨分析之间差别的一个细致研究仍然需要看哪个滤波器最适合特定的应用. 我们希望小波滤波器对极大重叠离散小波变换小波系数有一个更大的影响，因为这些更高地依赖于单独的小波滤波器的特性.

• **处理边界条件**

与离散小波变换一样，极大重叠离散小波变换利用了循环滤波，因此我们在4.11节所做的关于循环边界条件和镜像边界条件的简单选择对象的评论，可一样且很好地运用到极大重叠离散小波变换上. 接下来，我们考虑循环性影响极大重叠离散小波变换系数及多分辨分析的程度——我们的讨论与对离散小波变换的讨论是十分类似的，但相关的公式是十分不同的.

让我们先准确地确定哪些极大重叠离散小波变换系数被循环性假设所影响. 为此，仅需观察一下式(169a)中对$\widetilde{W}_{j,t}$和$\widetilde{V}_{j,t}$的定义，且了解t的哪些值能使我们在不改变这些系数定义的方式下，消除$X_{t-l \bmod N}$中的"求余"操作. 在形成每个系数时，指标范围从0到$L_j-1=(2^j-1)(L-1)$. 因为只要$t \geq L_j-1$，对所有的l就有

$t-l\bmod N=t-l$，所以被循环性所影响的极大重叠离散小波变换系数是

$$\widetilde{W}_{j,t} \quad 和 \quad \widetilde{V}_{j,t}, \quad t=0,\cdots,\min\{L_j-2,N-1\}. \tag{198a}$$

我们称上边的系数是极大重叠离散小波变换边界系数．如对$\mathbf{W}_j$和$\mathbf{V}_j$中的离散小波变换系数是正确的一样，边界系数全部产生在$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$的开始．因为这些向量的维数总是N，且L_j随着j的增大而增大，故边界系数的比例也随着j的增大而增大且当$L_j-1\geqslant N$时，即当$j\geqslant\log_2\left(\frac{N}{L-1}+1\right)$时，比为1.

让我们现在特殊化最接近对称的滤波器和coiflet小波滤波器，且假设$L_j\leqslant N$，以使$\min\{L_j-2,N-1\}=L_j-2$(即至少有一个系数不是边界系数)，4.8和4.9节的讨论表明，为得到与$\mathbf{X}$调准的滤波器输出，我们必须循环平移极大重叠离散小波变换的系数向量，而平移的量由最接近对称的滤波器或coiflet滤波器的相位特性指定．有了在式(114b)和(114a)中定义的$\nu_j^{(H)}$和$\nu_j^{(G)}$，这个步骤产生了循环平移向量$\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j$和$\mathcal{T}^{-|\nu_j^{(G)}|}\widetilde{\mathbf{V}}_j$，其元素可相对与$\mathbf{X}$相应元素有关联的相同实际次数而被绘制成图．因此，一些极大重叠离散小波变换边界系数将被转移到循环平移向量的末端．在$\mathcal{T}^{-|\nu_j^{(H)}|}\widetilde{\mathbf{W}}_j$中的边界小波系数指标由

$$t=0,\cdots,L_j-2-|\nu_j^{(H)}| \quad 及 \quad t=N-|\nu_j^{(H)}|,\cdots,N-1 \tag{198b}$$

给出；同样地，$\mathcal{T}^{-|\nu_j^{(G)}|}\widetilde{\mathbf{V}}_j$里的边界系数的指标由

$$t=0,\cdots,L_j-2-|\nu_j^{(G)}| \quad 及 \quad t=N-|\nu_j^{(G)}|,\cdots,N-1 \tag{198c}$$

给出(练习[5.11]将核对这些式是合理的，即我们总有$L_j-2\geqslant|\nu_j^{(H)}|$和$L_j-2\geqslant|\nu_j^{(G)}|$)．这些结果能够用来将垂线加在循环平移向量的图中，其解释为这些线描绘了除了循环性影响的那些系数外已绘制成图的那部分．作为例子，这些线作为图183的一部分出现，展示了心电图序列的LA(8)极大重叠离散小波变换.

最后，让我们研究一下循环性是如何影响极大重叠离散小波变换细节$\widetilde{\mathcal{D}}_j$和平滑$\widetilde{\mathcal{S}}_j$的(这一讨论应用到所有小波滤波器；即我们不再只限制到最接近对称或coiflet滤波器)．像式(172a)所表示的一样，细节$\widetilde{\mathcal{D}}_j$能通过用$\left\{\left|\widetilde{H}_j\left(\frac{k}{N}\right)\right|^2\right\}$给出的一个循环滤波器滤波$\mathbf{X}$，而从$\mathbf{X}$中直接得到．这个滤波器是通过滤波器$\{\tilde{h}_{j,l}\}$与本身互相关作用而形成的一个周期滤波器，即$\{\tilde{h}_{j,l}\}$的自相关(参见式(36e))．如果用$\{\tilde{h}_j\star\tilde{h}_{j,l}\}$表示这个滤波器，就可以用式(36c)写

$$\tilde{h}_j\star\tilde{h}_{j,l}=\sum_{u=-\infty}^{\infty}\tilde{h}_{j,u}\tilde{h}_{j,u+l}=\sum_{u=0}^{L_j-1}\tilde{h}_{j,u}\tilde{h}_{j,u+l}, \quad l=\cdots,-1,0,1,\cdots,$$

其中我们用到了事实：对于$l<0$和$l\geqslant L_j$，$\tilde{h}_{j,l}=0$(将上式与(36c)比较，回忆$\tilde{h}_{j,l}$是实值的，因此$\tilde{h}_{j,l}^*=\tilde{h}_{j,l}$)．从上面我们可以证明，当$|l|\geqslant L_j$时，$\tilde{h}_j\star\tilde{h}_{j,l}=0$；即

滤波器 $\{\tilde{h}_j \star \tilde{h}_{j,l}\}$ 的宽度 $2L_j-1$ 为奇数，并且以 $l=0$ 为中心. 因此，我们可以将 $\widetilde{\mathcal{D}}_j$ 的元素写为

$$\widetilde{\mathcal{D}}_{j,t} = \sum_{l=-(L_j-1)}^{L_j-1} \tilde{h}_j \star \tilde{h}_{j,l} X_{t-l \bmod N}, \quad t = 0,\cdots,N-1,$$

从此式可以明显地看到，只要我们有 $t-(L_j-1)\geqslant 0$ 和 $t+(L_j-1)\leqslant N-1$，元素 $\widetilde{\mathcal{D}}_{j,t}$ 就不受循环性的影响；即我们可以写

$$\widetilde{\mathcal{D}}_{j,t} = \sum_{l=-(L_j-1)}^{L_j-1} \tilde{h}_j \star \tilde{h}_{j,l} X_{t-l}, \quad t = L_j-1,\cdots,N-L_j.$$

$\widetilde{\mathcal{D}}_j$ 的那些受循环性影响的元素因此就是那些满足 $t=0$，…，L_j-2 或 $t=N-L_j+1$，…，$N-1$ 的指标. 同样的论证对 $\widetilde{\mathcal{S}}_j$ 也是有效的.

我们可以使用上面的结果将垂线放到 $\widetilde{\mathcal{D}}_j$ 和 $\widetilde{\mathcal{S}}_j$ 的图中以描绘出边界区域，即由循环性影响的区域. 这在图 183、184、186、192、194 和 195 中做过，并且也在图 196 做出的极大重叠离散小波变换平滑 $\widetilde{\mathcal{S}}_6$ 中(这些是底层的 4 个序列)做过. 从后面这个图可以看到，随着小波滤波器长度 L 的增加，边界区域的范围也增加. 对图 196，边界区域也被绘画在相应的离散小波变换平滑 $\mathcal{S}_6$ 上(上面绘出的 4 个序列). 对一给定的小波滤波器，$\mathcal{S}_6$ 和 $\widetilde{\mathcal{S}}_6$ 之间的比较表明，极大重叠离散小波变换边界区域比离散小波变换区域稍微大些，我们必须不可避免地为极大重叠离散小波变换多分辨分析的有吸引力的特性付出代价. 像 4.11 节所强调的，我们必须将边界区域看作是循环性影响的非常传统的度量标准——循环性对之有有害影响的区域通常比较小.

• *处理不是 2 的幂的抽样尺寸*

如同本章里很早就特别提示过的，极大重叠离散小波变换对所有的抽样尺寸都是有定义的. 和离散小波变换不一样，这就没有必要为特殊的适应来处理特定的抽样尺寸.

• *对极大重叠离散小波变换 J_0 层的选择*

我们讨论对于离散小波变换在选取 J_0 时，注意一个合理的选择主要依赖于眼前的时间序列——对极大重叠离散小波变换也是一样的. 然而，很有意思的是，如果 $N=2^J$，我们就不能使 J_0 比离散小波变换的 J 大(因为离散小波变换塔式算法当只剩下单个尺度系数时，就必须结束). 令 $J_0>J$ 在表面上将会提供一种获得关于时间序列在比序列的整个范围还大的尺度上变化的信息，这是与直觉上相反的. 练习[5.9c]表明，事实上如果令 $J_0>J$，就什么益处也得不到：当 $N=2^J$ 时，$\widetilde{\mathcal{S}}_{J_0}$ 的所有元素就必须与对所有 $J_0\geqslant J$ 的 $\overline{X}$ 的抽样均值相同. 因为式(172c)说 $\widetilde{\mathcal{D}}_{J_0+1}=\widetilde{\mathcal{S}}_{J_0}-\widetilde{\mathcal{S}}_{J_0+1}$，对 $J_0>J$ 层的所有极大重叠离散小波变换细节都得限制

为零．对不是 2 的幂的抽样尺寸，实际上 $\widetilde{\mathcal{D}}_{J_0}$ 随着 $2^{J_0}/N$ 的增大超过大约 1.5 时，快速地接近于零．

如果，诸如在离散小波变换讨论中那样，我们极力要求在$\widetilde{\mathbf{V}}_{J_0}$里至少有一个不是边界系数的极大重叠离散小波变换系数，经验法则要求条件 $J_0<\log_2\left(\frac{N}{L-1}+1\right)$；如果这个法则被认为是太传统了，那么可选择的上界是 $J_0\leqslant\log_2(N)$，或在极少数例子里，$J_0\leqslant\log_2(1.5N)$．这三个上界应该被仅仅认为是最初选择 J_0 的向导——J_0 的最后选择必须来自于相关应用的考虑因素．

5.12 小结

假设 $\mathbf{X}$ 是 N 维向量，它的元素是实值时间序列 $\{X_t: t=0,\cdots,N-1\}$，其中抽样尺寸 N 是任意的正整数．对于任意正整数 J_0，$\mathbf{X}$ 的 J_0 层极大重叠离散小波变换是一个由 J_0+1 个向量$\widetilde{\mathbf{W}}_1$，…，$\widetilde{\mathbf{W}}_{J_0}$ 和$\widetilde{\mathbf{V}}_{J_0}$ 组成的变换，所有这些向量都是 N 维的．向量$\widetilde{\mathbf{W}}_j$ 包含了与尺度 $\tau_j\equiv 2^{j-1}$ 上的变化有关的极大重叠离散小波变换小波系数，而$\widetilde{\mathbf{V}}_{J_0}$ 包含了与尺度 $\lambda_{J_0}\equiv 2^{J_0}$ 上的平均值有关的极大重叠离散小波变换尺度系数．在概念上，基于极大重叠离散小波变换系数的定义(见下面)，我们能写$\widetilde{\mathbf{W}}_j=\widetilde{\mathcal{W}}_j\mathbf{X}$ 和$\widetilde{\mathbf{V}}_{J_0}=\widetilde{\mathcal{V}}_{J_0}\mathbf{X}$，这里 $\widetilde{\mathcal{W}}_j$ 和 $\widetilde{\mathcal{V}}_{J_0}$ 是 $N\times N$ 阶矩阵(然而，实际上，极大重叠离散小波变换系数向量是通过下面的一个有效"塔式"算法所产生的)．时间序列 $\mathbf{X}$ 可以由它的极大重叠离散小波变换恢复：

$$\mathbf{X}=\sum_{j=1}^{J_0}\widetilde{\mathcal{W}}_j^{\mathrm{T}}\widetilde{\mathbf{W}}_j+\widetilde{\mathcal{V}}_{J_0}^{\mathrm{T}}\widetilde{\mathbf{V}}_{J_0}=\sum_{j=1}^{J_0}\widetilde{\mathcal{D}}_j+\widetilde{\mathcal{S}}_{J_0},$$

这借助于第 j 层极大重叠离散小波变换细节$\widetilde{\mathcal{D}}_j\equiv\widetilde{\mathcal{W}}_j^{\mathrm{T}}\widetilde{\mathbf{W}}_j$ 和 J_0 层极大重叠离散小波变换光滑$\widetilde{\mathcal{S}}_{J_0}\equiv\widetilde{\mathcal{V}}_{J_0}^{\mathrm{T}}\widetilde{\mathbf{V}}_{J_0}$，定义了一个基于极大重叠离散小波变换的 $\mathbf{X}$ 的多分辨分析．一个尺度对尺度的方差分析可以建立在能量分解

$$\|\mathbf{X}\|^2=\sum_{j=1}^{J_0}\|\widetilde{\mathbf{W}}_j\|^2+\|\widetilde{\mathbf{V}}_{J_0}\|^2$$

的基础上．和离散小波变换不一样，极大重叠离散小波变换不是一个规范正交变换，但它从来没有失去产生上面的多分辨分析和能量分解的能力，与离散小波变换产生的很类似．极大重叠离散小波变换的一个主要好处是，和离散小波变换不一样，它不苛刻地依赖于对一个时间序列的起点的假设．特别地，如果$\widetilde{\mathbf{W}}_1$，…，$\widetilde{\mathbf{W}}_{J_0}$ 和$\widetilde{\mathbf{V}}_{J_0}$ 是 $\mathbf{X}$ 的极大重叠离散小波变换，那么循环平移序列$\mathcal{T}^m\mathbf{X}$ 的极大重叠离散小波变换就是由$\mathcal{T}^m\widetilde{\mathbf{W}}_1$，…，$\mathcal{T}^m\widetilde{\mathbf{W}}_{J_0}$ 和$\mathcal{T}^m\widetilde{\mathbf{V}}_{J_0}$ 给出的．这一事实指出，$\mathcal{T}^m\mathbf{X}$ 的多分辨分析有细节$\mathcal{T}^m\widetilde{\mathcal{D}}_j$ 和光滑$\mathcal{T}^m\widetilde{\mathcal{S}}_{J_0}$；另一方面，由于 $\|\mathcal{T}^m\widetilde{\mathbf{W}}_j\|^2=\|\widetilde{\mathbf{W}}_j\|^2$ 和 $\|\mathcal{T}^m\widetilde{\mathbf{V}}_{J_0}\|^2=\|\widetilde{\mathbf{V}}_{J_0}\|^2$，能量分解仍然保持．

根据定义，$\widetilde{\mathbf{W}}_j$ 和$\widetilde{\mathbf{V}}_j$ 的元素是由滤波 $\mathbf{X}$ 得到的输出，即

$$\widetilde{W}_{j,t}\equiv\sum_{l=0}^{L_j-1}\tilde{h}_{j,l}X_{t-l\bmod N}\quad\text{和}\quad\widetilde{V}_{j,t}\equiv\sum_{l=0}^{L_j-1}\tilde{g}_{j,l}X_{t-l\bmod N},t=0,\cdots,N-1,$$

其中$\{\tilde{h}_{j,l}\}$和$\{\tilde{g}_{j,l}\}$是第 j 层极大重叠离散小波变换小波和尺度滤波器. 这些滤波器是根据对离散小波变换的第 j 层等价小波和尺度滤波器$\{h_{j,l}\}$和$\{g_{j,l}\}$按照$\tilde{h}_{j,l}=h_{j,l}/2^{j/2}$和$\tilde{g}_{j,l}=g_{j,l}/2^{j/2}$给出. 每个极大重叠离散小波变换滤波器均有宽度 $L_j\equiv(2^j-1)(L-1)+1$, 且能够由基本的极大重叠离散小波变换小波滤波器$\tilde{h}_{1,l}\equiv\tilde{h}_l\equiv h_l/\sqrt{2}$和与其相关的极大重叠离散小波变换尺度滤波器$\tilde{g}_{1,l}\equiv\tilde{g}_l\equiv(-1)^{l+1}\tilde{h}_{L-1-l}$产生. 具体地, 如果$\widetilde{H}(\cdot)$和$\widetilde{G}(\cdot)$分别是$\{\tilde{h}_l\}$和$\{\tilde{g}_l\}$的传递函数, 则$\{\tilde{h}_{j,l}\}$和$\{\tilde{g}_{j,l}\}$的传递函数就由

$$\widetilde{H}_j(f)\equiv\widetilde{H}(2^{j-1}f)\prod_{l=0}^{j-2}\widetilde{G}(2^l f)\quad \text{和} \quad \widetilde{G}_j(f)\equiv\prod_{l=0}^{j-1}\widetilde{G}(2^l f)$$

给出. 因为极大重叠离散小波变换滤波器仅仅是对离散小波变换相应的等价滤波器的重新尺度化形式, 故我们从前者离散小波变换得出的所有关系式经过明显调整后对后者极大重叠离散小波变换仍是有效的——最重要的在表 202 中给出. 如果令$\{\tilde{h}_{j,l}^{\circ}\}$和$\{\tilde{g}_{j,l}^{\circ}\}$是循环滤波器$\{\tilde{h}_{j,l}\}$和$\{\tilde{g}_{j,l}\}$以 N 为周期进行周期化后的形式, 那我们可以用对极大重叠离散小波变换系数的定义得到

$$\widetilde{W}_{j,t}=\sum_{l=0}^{N-1}\tilde{h}_{j,l}^{\circ}X_{t-l \bmod N}\quad \text{和} \quad \widetilde{V}_{j,t}=\sum_{l=0}^{N-1}\tilde{g}_{j,l}^{\circ}X_{t-l \bmod N}.$$

这些可以用来缩减矩阵$\widetilde{\mathcal{W}}_j$ 和$\widetilde{\mathcal{V}}_{J_0}$的元素. 特别地, 式(171c)表明了$\widetilde{\mathcal{W}}_j$ 的每一行是如何用$\{\tilde{h}_{j,l}^{\circ}\}$的 N 个元素来构造的(矩阵$\widetilde{\mathcal{V}}_{J_0}$是用$\{\tilde{g}_{j,l}^{\circ}\}$类似形成的).

在实际应用中, 矩阵$\widetilde{\mathcal{W}}_j$ 和$\widetilde{\mathcal{V}}_{J_0}$是不被明确使用的, 但更确切些, 组成极大重叠离散小波变换的向量是用"塔式"算法形成的, 而这一算法恰恰是使用基本极大重叠离散小波变换小波和尺度滤波器的. 在 j 层的极大重叠离散小波变换塔式算法是由$\widetilde{\mathbf{V}}_{j-1}\equiv[\widetilde{V}_{j-1,0}, \widetilde{V}_{j-1,1}, \cdots, \widetilde{V}_{j-1,N-1}]^T$ 的 N 个元素循环滤波构成的, 得到第 j 层极大重叠离散小波变换小波和尺度系数为

$$\widetilde{W}_{j,t}=\sum_{l=0}^{L-1}\tilde{h}_l\widetilde{V}_{j-1,t-2^{j-1}l \bmod N}\quad \text{和} \quad \widetilde{V}_{j,t}=\sum_{l=0}^{L-1}\tilde{g}_l\widetilde{V}_{j-1,t-2^{j-1}l \bmod N},\quad t=0,\cdots,N-1$$

(这些是$\widetilde{\mathbf{W}}_j$ 和$\widetilde{\mathbf{V}}_j$ 的元素). 由$\widetilde{\mathbf{V}}_0\equiv\mathbf{X}$, 我们从 $j=1$ 开始塔式算法, 且由 $j=2$, 3, $\cdots$, J_0 重复算法后, 得到所有的 J_0+1 层极大重叠离散小波变换系数向量(其他向量$\widetilde{\mathbf{V}}_1$, $\cdots$, $\widetilde{\mathbf{V}}_{J_0-1}$是尺度小于 λ_{J_0} 的极大重叠离散小波变换尺度系数, 它们是中间的计算量). 塔式算法可以用矩阵的记号表示为$\widetilde{\mathbf{W}}_j=\widetilde{\mathcal{B}}_j\widetilde{\mathbf{V}}_{j-1}$和$\widetilde{\mathbf{V}}_j=\widetilde{\mathcal{A}}_j\widetilde{\mathbf{V}}_{j-1}$, 其中$\widetilde{\mathcal{B}}_j$ 和$\widetilde{\mathcal{A}}_j$ 是 $N\times N$ 阶矩阵. 这些矩阵的元素是从长度为 N 的循环滤波器里取得的, 而这些滤波器是由周期化$\{\tilde{h}_l\}$和$\{\tilde{g}_l\}$的上抽样形式得到的. 特别地, 假设我们用因数 2^{j-1} 上抽样$\{\tilde{h}_l\}$(即在滤波器的每个元素间插入 $2^{j-1}-1$ 个零). 如果我们用$\{\tilde{h}_l^{(\uparrow 2^{j-1})\circ}\}$表示周期化这个上抽样长度为 N 的滤波器的结果, 那么$\widetilde{\mathcal{B}}_j$ 的第一行就由

$$[\tilde{h}_0^{(\uparrow 2^{j-1})\circ},\tilde{h}_{N-1}^{(\uparrow 2^{j-1})\circ},\tilde{h}_{N-2}^{(\uparrow 2^{j-1})\circ},\cdots,\tilde{h}_2^{(\uparrow 2^{j-1})\circ},\tilde{h}_1^{(\uparrow 2^{j-1})\circ}]$$

给出, 且剩下的 $N-1$ 行是将上边的分别向右循环平移 1, 2, $\cdots$, $N-1$ 个单位得到的.

表 202

$$\tilde h_l \equiv h_l/\sqrt{2} \qquad \tilde g_l \equiv g_l/\sqrt{2}$$

$$\{\tilde h_l\} \leftrightarrow \tilde H(\cdot) = \frac{1}{\sqrt{2}}H(\cdot) \qquad \{\tilde g_l\} \leftrightarrow \tilde G(\cdot) = \frac{1}{\sqrt{2}}G(\cdot)$$

$$\tilde h_l = (-1)^l \tilde g_{L-1-l} \qquad \tilde g_l \equiv (-1)^{l+1} \tilde h_{L-1-l}$$

$$\tilde H(f) = -e^{-i2\pi f(L-1)}\tilde G\left(\frac{1}{2}-f\right) \qquad \tilde G(f) = e^{-i2\pi f(L-1)}\tilde H\left(\frac{1}{2}-f\right)$$

$$\sum_l \tilde h_l = \tilde H(0) \equiv 0 \qquad \sum_l \tilde g_l = \tilde G(0) = 1$$

$$\sum_l \tilde h_l^2 \equiv \frac{1}{2} \qquad \sum_l \tilde g_l^2 = \frac{1}{2}$$

$$\sum_l \tilde h_l \tilde h_{l+2n} \equiv 0,\ n \neq 0 \qquad \sum_l \tilde g_l \tilde g_{l+2n} = 0,\ n \neq 0$$

$$\sum_l \tilde g_l \tilde h_{l+2n} = 0$$

$$\tilde{\mathcal{H}}(f) \equiv |\tilde H(f)|^2 = \frac{1}{2}\mathcal{H}(f) \qquad \tilde{\mathcal{G}}(f) \equiv |\tilde G(f)|^2 = \frac{1}{2}\mathcal{G}(f)$$

$$\tilde{\mathcal{H}}(f) + \tilde{\mathcal{H}}\left(f+\frac{1}{2}\right) = 1 \qquad \tilde{\mathcal{G}}(f) + \tilde{\mathcal{G}}\left(f+\frac{1}{2}\right) = 1$$

$$\tilde{\mathcal{G}}(f) + \tilde{\mathcal{H}}(f) = 1$$

$$\tilde W_{1,t} \equiv \sum_l \tilde h_l X_{t-l \bmod N} \qquad \tilde V_{1,t} \equiv \sum_l \tilde g_l X_{t-l \bmod N}$$

$$\tilde W_{j,t} = \sum_l \tilde h_l \tilde V_{j-1,t-2^{j-1}l \bmod N} \qquad \tilde V_{j,t} = \sum_l \tilde g_l \tilde V_{j-1,t-2^{j-1}l \bmod N}$$

$$\tilde h_{1,l} \equiv \tilde h_l,\ \tilde H_1(f) \equiv \tilde H(f) \qquad \tilde g_{1,l} \equiv \tilde g_l,\ \tilde G_1(f) \equiv \tilde G(f)$$

$$\tilde H_j(f) \equiv \tilde H(2^{j-1}f)\prod_{l=0}^{j-2}\tilde G(2^l f) \qquad \tilde G_j(f) \equiv \prod_{l=0}^{j-1}\tilde G(2^l f)$$

$$\tilde H_j(f) = \frac{1}{2^{j/2}}H_j(f) \qquad \tilde G_j(f) = \frac{1}{2^{j/2}}G_j(f)$$

$$\tilde H_j(f) = \tilde H(2^{j-1}f)\tilde G_{j-1}(f) \qquad \tilde G_j(f) = \tilde G(2^{j-1}f)\tilde G_{j-1}(f)$$

$$\{\tilde h_{j,l}\} \leftrightarrow \tilde H_j(\cdot) \qquad \{\tilde g_{j,l}\} \leftrightarrow \tilde G_j(\cdot)$$

$$\sum_l \tilde h_{j,l} = \tilde H_j(0) = 0 \qquad \sum_l \tilde g_{j,l} = \tilde G_j(0) = 1$$

$$\sum_l \tilde h_{j,l}^2 = \frac{1}{2^j} \qquad \sum_l \tilde g_{j,l}^2 = \frac{1}{2^j}$$

$$\sum_l \tilde h_{j,l}\tilde h_{j,l+2^j n} = 0,\ n \neq 0 \qquad \sum_l \tilde g_{j,l}\tilde g_{j,l+2^j n} = 0,\ n \neq 0$$

$$\sum_l \tilde g_{j,l}\tilde h_{j,l+2^j n} = 0$$

$$\tilde{\mathcal{H}}_j(f) \equiv |\tilde H_j(f)|^2 = \frac{1}{2^j}\mathcal{H}_j(f) \qquad \tilde{\mathcal{G}}_j(f) \equiv |\tilde G_j(f)|^2 = \frac{1}{2^j}\mathcal{G}_j(f)$$

$$\tilde W_{j,t} \equiv \sum_l \tilde h_{j,l} X_{t-l \bmod N} \qquad \tilde V_{j,t} \equiv \sum_l \tilde g_{j,l} X_{t-l \bmod N}$$

注：涉及极大重叠离散小波变换的小波和尺度滤波器的主要关系式（对这里用到的 l 之和的极限与表 154 中的约定一样）.

给定$\widetilde{\mathbf{W}}_j$和$\widetilde{\mathbf{V}}_j$，我们能够经由极大重叠离散小波逆变换塔式算法的第j步恢复$\widetilde{\mathbf{V}}_{j-1}$的元素，即

$$\widetilde{V}_{j-1,t} = \sum_{l=0}^{L-1} \tilde{h}_l \widetilde{W}_{j,t+2^{j-1}l \bmod N} + \sum_{l=0}^{L-1} \tilde{g}_l \widetilde{V}_{j,t+2^{j-1}l \bmod N}, \quad t=0,1,\cdots,N-1.$$

上式可用矩阵的记号表示为

$$\widetilde{\mathbf{V}}_{j-1} = \widetilde{\mathcal{B}}_j^{\mathrm{T}} \widetilde{\mathbf{W}}_j + \widetilde{\mathcal{A}}_j^{\mathrm{T}} \widetilde{\mathbf{V}}_j.$$

如果我们由$\widetilde{\mathbf{V}}_0=\mathbf{X}$出发且结合上式对$j=1,\cdots,J_0$的结果，就会得到

$$\mathbf{X} = \widetilde{\mathcal{B}}_1^{\mathrm{T}} \widetilde{\mathbf{W}}_1 + \widetilde{\mathcal{A}}_1^{\mathrm{T}} \widetilde{\mathcal{B}}_2^{\mathrm{T}} \widetilde{\mathbf{W}}_2 + \widetilde{\mathcal{A}}_1^{\mathrm{T}} \widetilde{\mathcal{A}}_2^{\mathrm{T}} \widetilde{\mathcal{B}}_3^{\mathrm{T}} \widetilde{\mathbf{W}}_3 + \cdots + \widetilde{\mathcal{A}}_1^{\mathrm{T}} \cdots \widetilde{\mathcal{A}}_{J_0-1}^{\mathrm{T}} \widetilde{\mathcal{B}}_{J_0}^{\mathrm{T}} \widetilde{\mathbf{W}}_{J_0}$$

$$+ \widetilde{\mathcal{A}}_1^{\mathrm{T}} \cdots \widetilde{\mathcal{A}}_{J_0-1}^{\mathrm{T}} \widetilde{\mathcal{A}}_{J_0}^{\mathrm{T}} \widetilde{\mathbf{V}}_{J_0}.$$

从上式，我们能以$\widetilde{\mathcal{D}}_j=\widetilde{\mathcal{A}}_1^{\mathrm{T}}\cdots\widetilde{\mathcal{A}}_{j-1}^{\mathrm{T}}\widetilde{\mathcal{B}}_j^{\mathrm{T}}\widetilde{\mathbf{W}}_j$和$\widetilde{\mathcal{S}}_{J_0}=\widetilde{\mathcal{A}}_1^{\mathrm{T}}\cdots\widetilde{\mathcal{A}}_{J_0-1}^{\mathrm{T}}\widetilde{\mathcal{A}}_{J_0}^{\mathrm{T}}\widetilde{\mathbf{V}}_{J_0}$确定极大重叠离散小波变换细节和光滑的表示形式. 因为$\widetilde{\mathcal{D}}_j=\widetilde{\mathcal{W}}_j^{\mathrm{T}}\widetilde{\mathbf{W}}_j$和$\widetilde{\mathcal{S}}_{J_0}=\widetilde{\mathcal{V}}_{J_0}^{\mathrm{T}}\widetilde{\mathbf{V}}_{J_0}$，我们得到$\widetilde{\mathcal{W}}_j=\widetilde{\mathcal{B}}_j\widetilde{\mathcal{A}}_{j-1}\cdots\widetilde{\mathcal{A}}_1$和$\widetilde{\mathcal{V}}_j=\widetilde{\mathcal{A}}_j\widetilde{\mathcal{A}}_{j-1}\cdots\widetilde{\mathcal{A}}_1$.

给定$\widetilde{\mathcal{W}}_j$和$\widetilde{\mathcal{V}}_{J_0}$的结构，我们能明确地表示$\widetilde{\mathcal{D}}_j$和$\widetilde{\mathcal{S}}_{J_0}$的元素

$$\widetilde{\mathcal{D}}_{j,t} = \sum_{l=0}^{N-1} \tilde{h}_{j,l}^{\circ} \widetilde{W}_{j,t+l \bmod N} \quad 和 \quad \widetilde{\mathcal{S}}_{J_0,t} = \sum_{l=0}^{N-1} \tilde{g}_{J_0,l}^{\circ} \widetilde{V}_{J_0,t+l \bmod N}, t=0,1,\cdots,N-1.$$

通过结合上式与包含$\tilde{h}_{j,l}^{\circ}$和$\tilde{g}_{J_0,l}^{\circ}$的对$\widetilde{W}_{j,t}$和$\widetilde{V}_{J_0,t}$的表达式，我们可证明$\widetilde{\mathcal{D}}_j$和$\widetilde{\mathcal{S}}_{J_0}$能通过传递函数由$|\widetilde{H}_j(\cdot)|^2$和$|\widetilde{G}_{J_0}(\cdot)|^2$给出的滤波器，对时间序列$\mathbf{X}$进行滤波而直接得到. 这些滤波器的相位函数都恒为零，故基于极大重叠离散小波变换的多分辨分析的每个成分都是与零相位滤波器相关的. 零相位特性是很重要的，因为它允许我们用$\mathbf{X}$里的结果有意义地排列细节和光滑里的结果.

最后，如果对$\mathbf{X}$的抽样尺寸是为我们能计算J_0层离散小波变换，我们注意到下列对任意$j\leqslant J_0$的第j层离散小波变换和极大重叠离散小波变换小波系数之间的关系式：

$$W_{j,t} = 2^{j/2} \widetilde{W}_{j,2^j(t+1)-1}, \quad t=0,\cdots,N_j-1,$$

其中$N_j\equiv N/2^j$（上式是式(96d)). 类似地，尺度系数有关系$V_{J_0,t}=2^{J_0/2}\widetilde{V}_{J_0,2^{J_0}(t+1)-1}$.

5.13 练习

[5.1] 假设 $\mathbf{X}$ 包含偶数抽样尺寸 N 的一个时间序列. 令 $\mathcal{D}_1$ 和 $\mathcal{S}_1$ 是以 $J_0=1$ 层部分离散小波变换为基础的 $\mathbf{X}$ 的一个多分辨分析的细节和光滑,因此我们有 $\mathbf{X}=\mathcal{D}_1+\mathcal{S}_1$. 令 $\mathcal{D}_{\mathcal{T},1}$ 和 $\mathcal{S}_{\mathcal{T},1}$ 是对 $\mathcal{T}\mathbf{X}$ 的一个多分辨分析的相应的量(即 $\mathbf{X}$ 向右一个单位的循环平移像式(52a)). 因此我们有 $\mathcal{T}\mathbf{X}=\mathcal{D}_{\mathcal{T},1}+\mathcal{S}_{\mathcal{T},1}$,且因此 $\mathbf{X}=\mathcal{T}^{-1}\mathcal{D}_{\mathcal{T},1}+\mathcal{T}^{-1}\mathcal{S}_{\mathcal{T},1}$. 假设 $\widetilde{\mathcal{D}}_1$ 和 $\widetilde{\mathcal{S}}_1$ 是基于 $J_0=1$ 层极大重叠离散小波变换的 $\mathbf{X}$ 的多分辨分析的细节和光滑,因此有 $\mathbf{X}=\widetilde{\mathcal{D}}_1+\widetilde{\mathcal{S}}_1$. 证明

$$\widetilde{\mathcal{D}}_1=\frac{1}{2}(\mathcal{D}_1+\mathcal{T}^{-1}\mathcal{D}_{\mathcal{T},1}) \quad \text{和} \quad \widetilde{\mathcal{S}}_1=\frac{1}{2}(\mathcal{S}_1+\mathcal{T}^{-1}\mathcal{S}_{\mathcal{T},1}).$$

证明我们也可以写为

$$\widetilde{\mathcal{D}}_1=\frac{1}{N}\sum_{n=0}^{N-1}\mathcal{T}^{-n}\mathcal{D}_{\mathcal{T}^n,1} \quad \text{和} \quad \widetilde{\mathcal{S}}_1=\frac{1}{N}\sum_{n=0}^{N-1}\mathcal{T}^{-n}\mathcal{S}_{\mathcal{T}^n,1}.$$

其中 $\mathcal{D}_{\mathcal{T}^n,1}$ 和 $\mathcal{S}_{\mathcal{T}^n,1}$ 构成对 $\mathcal{T}^n\mathbf{X}$ 的基于离散小波变换的多分辨分析.(极大重叠离散小波变换细节和光滑作为对 $\mathbf{X}$ 平移形式的平移的离散小波变换细节和光滑的平均值,这一解释一般也是有效的,因为当 $N=2^J$ 和 $1\leqslant j\leqslant J_0\leqslant J$ 时,我们有

$$\widetilde{\mathcal{D}}_j=\frac{1}{2^j}\sum_{n=0}^{2^j-1}\mathcal{T}^{-n}\mathcal{D}_{\mathcal{T}^n,j} \quad \text{和} \quad \widetilde{\mathcal{S}}_{J_0}=\frac{1}{2^{J_0}}\sum_{n=0}^{2^{J_0}-1}\mathcal{T}^{-n}\mathcal{S}_{\mathcal{T}^n,J_0}, \tag{204}$$

还有

$$\widetilde{\mathcal{D}}_j=\frac{1}{N}\sum_{n=0}^{N-1}\mathcal{T}^{-n}\mathcal{D}_{\mathcal{T}^n,j} \quad \text{和} \quad \widetilde{\mathcal{S}}_{J_0}=\frac{1}{N}\sum_{n=0}^{N-1}\mathcal{T}^{-n}\mathcal{S}_{\mathcal{T}^n,J_0}.$$

对于时间序列各种平移形成的离散小波变换概念,然后平均基于这些离散小波变换的适当平移过的综合,这是在 Coifman and Donoho,1995,讨论过的"循环旋转"的主要概念.)

[5.2] 证明 $\widetilde{\mathcal{P}}_1^T$ 是对 $\widetilde{\mathcal{P}}_1$ 的 Moore-Penrose 广义逆(Peterson,1996).

提示:你可以用这样的结果,如果 $\widetilde{\mathcal{M}}$ 是一个 $N\times 2N$ 阶矩阵满足 $\widetilde{\mathcal{M}}\widetilde{\mathcal{P}}_1=I_N$,并且 $\widetilde{\mathcal{P}}_1\widetilde{\mathcal{M}}$ 是对称的,那么 $\widetilde{\mathcal{M}}$ 就是对 $\widetilde{\mathcal{P}}_1$ 的 Moore-Penrose 的逆(Rao and Mitra,1971).

[5.3] 式(176)表示了对 $N=12$ 和 $L=4$ 的 $\widetilde{\mathcal{B}}_2$. 对这种情形 $\widetilde{\mathcal{B}}_3$ 看起来会是什么样呢?

[5.4] 利用5.5节的评论和扩展[1]中的伪代码,用你最喜欢的计算机语言实现极大重叠离散小波变换塔式算法和它的逆. 用D(4)小波和尺度滤波器,计算并做出(或列出值)16个点的时间序列 $\{X_{1,t}\}$ 的 $\widetilde{\mathbf{W}}_j$ 和 $\widetilde{\mathbf{V}}_j$ ($j=1,2$ 和 3),参见对图42的说明. 通过表明你能从 $\widetilde{\mathbf{W}}_j$ 和 $\widetilde{\mathbf{V}}_j$ ($j=1,2$ 和 3)重构 $\widetilde{\mathbf{V}}_{j-1}$ 来验证,你的极大重叠离散小波逆变换塔式算法的实现是正确的(这里我们定义 $\widetilde{\mathbf{V}}_0=\mathbf{X}$). 验证,通过适当子抽样和重尺度化 $\widetilde{\mathbf{W}}_j$ 和 $\widetilde{\mathbf{V}}_j$,你就能得到在练习[4.17]计算出的离散小波变换 $\mathbf{W}_j$ 和 $\mathbf{V}_j$.

[5.5] 对练习[4.19]中考虑的同样的时间序列 X_t,用下列 4 个小波:哈尔、D(4)、C(6)和 LA(8)中的任何两个计算并绘出 $k=1$ 到 6 的极大重叠离散小波变换细节 $\widetilde{\mathcal{D}}_k$. 解释与(a){$X_t$}是怎样被定义的(b)用到的特定小波,相关的细节序列的形式.

[5.6] 对列在图 42 标题上的 16 个点的时间序列 $\mathbf{X}=[X_{1,0},X_{1,1},\cdots,X_{1,15}]^T$,以哈尔离散小波变换为基础计算 $J_0=2$ 层的多分辨分析,并且对循环平移序列 $\mathcal{T}^n\mathbf{X}$, $n=1,2$ 和 3,计算出相应的多分辨分析. 假设 $\mathcal{D}_{\mathcal{T}^n,1}$、$\mathcal{D}_{\mathcal{T}^n,2}$ 和 $\mathcal{S}_{\mathcal{T}^n,2}$ 表示 $\mathcal{T}^n\mathbf{X}$, $n=0,\cdots,3$ 的多分辨分析的细节和光滑. 在数值上证明方程(204)成立.

[5.7] 证明对所有的 j, $\|\widetilde{\mathcal{D}}_j\|^2\leqslant\|\widetilde{\mathbf{W}}_j\|^2$ 和 $\|\widetilde{\mathcal{S}}_j\|^2\leqslant\|\widetilde{\mathbf{V}}_j\|^2$ 成立.

[5.8] 证明对所有的 j, $\widetilde{\mathcal{B}}_j^T\widetilde{\mathcal{B}}_j+\widetilde{\mathcal{A}}_j^T\widetilde{\mathcal{A}}_j=I_N$ 成立(这类结果有时被称为是"恒等分解").

[5.9] 这里我们考虑习题[171b]的三个部分.

(a)证明,如果抽样尺寸是 2 的幂,即 $N=2^J$,且选择 $J_0\geqslant J$,那么 $\widetilde{\mathbf{V}}_{J_0}$ 里所有的极大重叠离散小波变换尺度系数事实上与 $\overline{X}$ 相等,然而,如果 N 不是 2 的幂,那么在 $\widetilde{\mathbf{V}}_{J_0}$ 的元素中就有一些变化. 我们可用这个来证明式(171a)可以缩减为

$$\hat{\sigma}_X^2=\frac{1}{N}\sum_{j=1}^{J_0}\|\widetilde{\mathbf{W}}_j\|^2,\quad 当\ N=2^J\quad 和\quad J_0\geqslant J.$$

(b)证明,对任何抽样尺寸 N 和任何层 J_0,极大重叠离散小波变换平滑 $\widetilde{\mathcal{S}}_{J_0}$ 中元素的抽样均值与 $\overline{X}$ 相等,而任何极大重叠离散小波变换细节 $\widetilde{\mathcal{D}}_j$ 的抽样均值对任意 j 层来说均为零.

(c)证明,如果 $N=2^J$ 且我们选择 $J_0\geqslant J$,那么 $\widetilde{\mathcal{S}}_{J_0}$ 的所有元素事实上是与 $\overline{X}$ 相同的,然而,如果 N 不是 2 的幂,那么在 $\widetilde{\mathcal{S}}_{J_0}$ 的元素中就会有一些变化.

[5.10] 证明,对哈尔小波滤波器,我们有 $\widetilde{H}(f)+\widetilde{G}(f)=1$. 证明这个恒等分解能被用来形成一个加性分解,并且是直接建立在极大重叠离散小波变换哈尔小波系数基础上的(C. A. Greenhall,个人评论). 这种方法与通常的哈尔极大重叠离散小波变换的多分辨分析相比较有什么样的好处和坏处呢?

[5.11] 证明,对在 4.8 节和 4.9 节分别讨论的所有最接近对称和 coiflet 小波滤波器,我们总有 $L_j-2\geqslant|\nu_j^{(H)}|$ 和 $L_j-2\geqslant|\nu_j^{(G)}|$,这里 $|\nu_j^{(H)}|$ 和 $|\nu_j^{(G)}|$ 是在式(114b)和(114a)里定义的(这个练习证明了在式(198b)和(198c)里给定的 t 的极限是有意义的).

第 6 章

离散小波包变换

6.0 引言

在第 4 章里，我们讨论了离散小波变换，其实质上是将一个时间序列 $\mathbf{X}$ 分解成与不同尺度和时间有关的系数. 因此，我们可以认为 $\mathbf{X}$ 的离散小波变换是一个"时间/尺度"的分解. 对一个给定的尺度 $\tau_j=2^{j-1}$，其小波系数告诉我们，$\mathbf{X}$ 的局部加权平均值怎样从一个平均周期到下一个周期的. 尺度 τ_j 给出了加权平均值的时间上的有效宽度(即局部化等级). 因为离散小波变换可以借助于滤波器系统的阐述，故我们可以将尺度概念和某些频带联系起来. 对于尺度 τ_j 产生小波系数的等价滤波器近似是一个通带由$[1/2^{j+1}, 1/2^j]$给出的带通滤波器. 对于抽样尺寸 $N=2^J$，$N-1$ 个小波系数组成——当结合在一起时——频率区间$[1/2^{J+1}, 1/2]$的一个倍频程带分解，而单个的尺度系数与区间$[0, 1/2^{J+1}]$相关. 总之，离散小波变换系数将频率区间$[0, 1/2]$分解成邻接的单个区间.

本章中，我们将考虑离散小波包变换(Discrete Wavelet Packet Transform，简称 DWPT)，它能作为是规范正交变换集中的任何一个，其中的每一个都能通过对离散小波变换塔式算法的一个简单变形计算出来. 每个离散小波包变换都与 j 层有关，而且第 j 层离散小波包变换将频率区间$[0, 1/2]$分解成 2^j 个相等的单个区间. 因为第 $J-1$层分解将频率区间$[0, 1/2]$分成 $N/2=2^{J-1}$个相等的区间，因此，有一个模仿的离散小波包变换——但与之不相同——由离散傅里叶变换给出的$[0, 1/2]$上的分解. 当$j=1, \cdots, J-1$ 时，相应的离散小波包变换产生一个称为"时间/频率"的分解，因为每个离散小波包变换系数可以局部化为特殊的频带和特殊的时间区间(这在本质上与所谓的短时傅里叶变换相似，而短时傅里叶变换实质上是通过将离散傅里叶变换应用到从 $\mathbf{X}$ 中提取的子序列上来计算的). 我们也可以通过将从不同离散小波包变换仔细筛选出的基向量集合起来产生一个更大的规范正交变换集. 事实上，离散小波变换和所有部分离散小波变换都能够用不同离散小波包变换里的基向量形成，因此这个方案导向了一个在时间/尺度和时间/频率分解之间充当桥梁的灵活的变换集.

本章余下的部分是如下组织的. 首先，在6.1节中讨论离散小波包变换的基本概念，和离散小波变换一样，尤其强调了这个变换只依赖于小波滤波器$\{h_l\}$及与之相关的尺度滤波器$\{g_l\}$. 在6.2节中我们将给出利用太阳物理时间序列的离散小波包变换分析的一个例子. 因为离散小波包变换能用来产生规范正交变换的整个集合，所以在6.3节里我们考虑了对一个给定的时间序列定义一个"最优"变换的方法，然后在6.4节说明这种方法在太阳物理序列当中的用途. 当用Daubechies最接近对称或coiflet滤波器产生一个离散小波包变换时，我们能得到被近似认为是一个零相位滤波器输出的变换系数. 在6.5节中，我们将讨论需要得到这种近似零相位特性的时间平移. 在6.6节中，我们定义与极大重叠离散小波变换相似的极大重叠离散小波包变换(Maximal Overlap Discrete Wavelet Packet Transform，简称MODWPT)，且在6.7节中用太阳物理序列提供了极大重叠离散小波包变换分析的一个例子. 接着将在6.8节讨论"匹配追踪"的概念，它允许我们通过用从一个大的向量集(包括通过合并不同的离散小波包变换或极大重叠离散小波包变换形成的向量)里选出的少数向量的线性组合来逼近时间序列. 在6.9节中将匹配追踪算法的概念应用到了新奥尔良市的子潮海平面序列. 最后，在6.10和6.11节中，我们将给出本章的内容小结和一些练习.

6.1 基本概念

像第4章中，令$\mathbf{W}=\mathcal{W}\mathbf{X}$表示用$N\times N$阶正交离散小波变换矩阵$\mathcal{W}$变换$\mathbf{X}$得到的小波系数(为方便起见，我们假设对某个整数$J$，$N=2^J$). 实际上，矩阵$\mathcal{W}$实际上没有产生，它隐含在塔式算法中. 这个算法的第一步可以用矩阵的符号描述为

$$\mathcal{P}_1\mathbf{X}=\begin{bmatrix}\mathcal{B}_1\\ \mathcal{A}_1\end{bmatrix}\mathbf{X}=\begin{bmatrix}\mathbf{W}_1\\ \mathbf{V}_1\end{bmatrix}\equiv\begin{bmatrix}\mathbf{W}_{1,1}\\ \mathbf{W}_{1,0}\end{bmatrix},$$

其中$\mathbf{W}_{1,1}\equiv\mathbf{W}_1$且$\mathbf{W}_{1,0}\equiv\mathbf{V}_1$. 矩阵$\mathcal{P}_1$是规范正交阵，因为

$$\mathcal{P}_1\mathcal{P}_1^T=\begin{bmatrix}\mathcal{B}_1\\ \mathcal{A}_1\end{bmatrix}[\mathcal{B}_1^T,\mathcal{A}_1^T]=\begin{bmatrix}\mathcal{B}_1\mathcal{B}_1^T & \mathcal{B}_1\mathcal{A}_1^T\\ \mathcal{A}_1\mathcal{B}_1^T & \mathcal{A}_1\mathcal{A}_1^T\end{bmatrix}=\begin{bmatrix}I_{\frac{N}{2}} & 0_{\frac{N}{2}}\\ 0_{\frac{N}{2}} & I_{\frac{N}{2}}\end{bmatrix}=I_N,$$

这由于我们已经构造了$\frac{N}{2}\times N$阶矩阵$\mathcal{B}_1$和$\mathcal{A}_1$使得

$$\mathcal{B}_1\mathcal{B}_1^T=\mathcal{A}_1\mathcal{A}_1^T=I_{\frac{N}{2}}\quad 和\quad \mathcal{B}_1\mathcal{A}_1^T=\mathcal{A}_1\mathcal{B}_1^T=0_{\frac{N}{2}}$$

(这里$I_{\frac{N}{2}}$是$\frac{N}{2}\times\frac{N}{2}$阶单位矩阵，而$0_{\frac{N}{2}}$表示一个所有元素均为0的$\frac{N}{2}\times\frac{N}{2}$阶矩阵). 回忆一下，对一定的逼近阶，$\mathbf{W}_{1,1}$中的小波系数表示了在$\mathbf{X}$里频率$f$满足$|f|\in\left[\frac{1}{4},\frac{1}{2}\right]$的频率内容，而$\mathbf{W}_{1,0}$表示了$|f|\in\left[0,\frac{1}{4}\right]$时的频率含量. 变换$\mathcal{P}_1$对应于$J_0=1$层的部分离散小波变换(参见4.7节).

类似地，在塔式算法第二步的末尾，我们有一个变换($J_0=2$层的部分离散小波变换)表示为

$$\begin{bmatrix}\mathcal{B}_1\\ \mathcal{B}_2\mathcal{A}_1\\ \mathcal{A}_2\mathcal{A}_1\end{bmatrix}\mathbf{X}=\begin{bmatrix}\mathbf{W}_1\\ \mathbf{W}_2\\ \mathbf{V}_2\end{bmatrix}\equiv\begin{bmatrix}\mathbf{W}_{1,1}\\ \mathbf{W}_{2,1}\\ \mathbf{W}_{2,0}\end{bmatrix},$$

其中，和以前一样 $\mathbf{W}_{1,1}\equiv\mathbf{W}_1$，而 $\mathbf{W}_{2,1}\equiv\mathbf{W}_2$ 和 $\mathbf{W}_{2,0}\equiv\mathbf{V}_2$. 由构造，我们有

$$\mathcal{B}_2\mathcal{B}_2^T=\mathcal{A}_2\mathcal{A}_2^T=I_{\frac{N}{4}} \text{ 和 } \mathcal{B}_2\mathcal{A}_2^T=\mathcal{A}_2\mathcal{B}_2^T=0_{\frac{N}{4}},$$

因此很容易地得出此变换是规范正交的，因为

$$\begin{bmatrix}\mathcal{B}_1\\ \mathcal{B}_2\mathcal{A}_1\\ \mathcal{A}_2\mathcal{A}_1\end{bmatrix}[\mathcal{B}_1^T,\mathcal{A}_1^T\mathcal{B}_2^T,\mathcal{A}_1^T\mathcal{A}_2^T]=\begin{bmatrix}\mathcal{B}_1\mathcal{B}_1^T & \mathcal{B}_1\mathcal{A}_1^T\mathcal{B}_2^T & \mathcal{B}_1\mathcal{A}_1^T\mathcal{A}_2^T\\ \mathcal{B}_2\mathcal{A}_1\mathcal{B}_1^T & \mathcal{B}_2\mathcal{A}_1\mathcal{A}_1^T\mathcal{B}_2^T & \mathcal{B}_2\mathcal{A}_1\mathcal{A}_1^T\mathcal{A}_2^T\\ \mathcal{A}_2\mathcal{A}_1\mathcal{B}_1^T & \mathcal{A}_2\mathcal{A}_1\mathcal{A}_1^T\mathcal{B}_2^T & \mathcal{A}_2\mathcal{A}_1\mathcal{A}_1^T\mathcal{A}_2^T\end{bmatrix}$$

$$=\begin{bmatrix}I_{\frac{N}{2}} & 0_{\frac{N}{2}}\mathcal{B}_2^T & 0_{\frac{N}{2}}\mathcal{A}_2^T\\ \mathcal{B}_2 0_{\frac{N}{2}} & \mathcal{B}_2\mathcal{B}_2^T & \mathcal{B}_2\mathcal{A}_2^T\\ \mathcal{A}_2 0_{\frac{N}{2}} & \mathcal{A}_2\mathcal{B}_2^T & \mathcal{A}_2\mathcal{A}_2^T\end{bmatrix}$$

$$=\begin{bmatrix}I_{\frac{N}{2}} & 0_{\frac{N}{2},\frac{N}{4}} & 0_{\frac{N}{2},\frac{N}{4}}\\ 0_{\frac{N}{4},\frac{N}{2}} & I_{\frac{N}{4}} & 0_{\frac{N}{4}}\\ 0_{\frac{N}{4},\frac{N}{2}} & 0_{\frac{N}{4}} & I_{\frac{N}{4}}\end{bmatrix}=I_N,$$

其中 $0_{\frac{N}{2},\frac{N}{4}}$ 是一个 $\frac{N}{2}\times\frac{N}{4}$ 阶零矩阵. 借助于频率域，$\mathbf{W}_{1,1}$ 与频率区间 $\left[\frac{1}{4},\ \frac{1}{2}\right]$ 有关；$\mathbf{W}_{2,1}$ 与区间 $\left[\frac{1}{8},\ \frac{1}{4}\right]$ 有关；$\mathbf{W}_{2,0}$ 与区间 $\left[0,\ \frac{1}{8}\right]$ 有关.

注意到第二步的规范正交变换是通过使 $\mathcal{B}_1$ 里的向量不变，用 $\mathcal{B}_2$ 和 $\mathcal{A}_2$"旋转" $\mathcal{A}_1$ 里的向量，从第一步里得到的. 相反，如果我们让 $\mathcal{A}_1$ 里的向量不变，旋转 $\mathcal{B}_1$ 里的向量，会发生什么呢？那么，我们将会有一个系数由

$$\begin{bmatrix}\mathbf{W}_{2,3}\\ \mathbf{W}_{2,2}\\ \mathbf{W}_{1,0}\end{bmatrix}\equiv\begin{bmatrix}\mathcal{A}_2\mathcal{B}_1\\ \mathcal{B}_2\mathcal{B}_1\\ \mathcal{A}_1\end{bmatrix}\mathbf{X}$$

给出的变换.

练习[208] 证明上面的变换是规范正交的. ◁

进而，假如我们把所有的第二层的结果结合在一起：

$$\begin{bmatrix}\mathbf{W}_{2,3}\\ \mathbf{W}_{2,2}\\ \mathbf{W}_{2,1}\\ \mathbf{W}_{2,0}\end{bmatrix}=\begin{bmatrix}\mathcal{A}_2\mathcal{B}_1\\ \mathcal{B}_2\mathcal{B}_1\\ \mathcal{B}_2\mathcal{A}_1\\ \mathcal{A}_2\mathcal{A}_1\end{bmatrix}\mathbf{X}. \tag{208a}$$

练习[6.1]将会验证这个变换也是规范正交的. 正交性推出 $\mathbf{X}$ 的一个加性分解

$$\mathbf{X}=[\mathcal{B}_1^T\mathcal{A}_2^T,\mathcal{B}_1^T\mathcal{B}_2^T,\mathcal{A}_1^T\mathcal{B}_2^T,\mathcal{A}_1^T\mathcal{A}_2^T]\begin{bmatrix}\mathbf{W}_{2,3}\\ \mathbf{W}_{2,2}\\ \mathbf{W}_{2,1}\\ \mathbf{W}_{2,0}\end{bmatrix} \tag{208b}$$

$$=\mathcal{B}_1^T\mathcal{A}_2^T\mathbf{W}_{2,3}+\mathcal{B}_1^T\mathcal{B}_2^T\mathbf{W}_{2,2}+\mathcal{A}_1^T\mathcal{B}_2^T\mathbf{W}_{2,1}+\mathcal{A}_1^T\mathcal{A}_2^T\mathbf{W}_{2,0},$$

它是仿制通常离散小波变换多分辨分析的；同样，我们有

$$\|\mathbf{X}\|^2=\|\mathbf{W}_{2,0}\|^2+\|\mathbf{W}_{2,1}\|^2+\|\mathbf{W}_{2,2}\|^2+\|\mathbf{W}_{2,3}\|^2,$$

由此我们可以产生一个与建立在离散小波变换基础上很类似的方差分析(ANOVA). 图 210a描述了 $\mathbf{X}$ 到 $\mathbf{W}_{2,0}$、$\mathbf{W}_{2,1}$、$\mathbf{W}_{2,2}$和 $\mathbf{W}_{2,3}$的分析.

在式(208a)和图 210a 中，$\mathbf{W}_{2,n}$的序是尤其有用的，因为 n 对应着频率顺序：$\mathbf{W}_{2,0}$与标称频率区间$\left[0, \frac{1}{8}\right]$有关；$\mathbf{W}_{2,1}$与$\left[\frac{1}{8}, \frac{1}{4}\right]$有关；$\mathbf{W}_{2,2}$与$\left[\frac{1}{4}, \frac{3}{8}\right]$有关；$\mathbf{W}_{2,3}$与$\left[\frac{3}{8}, \frac{1}{2}\right]$有关. 这样的一个顺序在 Wickerhauser(1994)中称为连续顺序. 注意图 210a 中表示的第二步滤波分别是对 $\mathbf{W}_{2,n}$，$n=0, \cdots, 3$，的低通、高通、高通和低通滤波. 对这种低通和高通滤波顺器序的解释在 4.4 节里就给出了，这里我们注意到下面两个重要的事实：

- 一方面，用尺度滤波器$\{g_l\}$循环滤波 $\mathbf{X}$，然后用下 2 抽样得到序列 $\mathbf{V}_1=\mathbf{W}_{1,0}$，它在$f\in[0, 1/2]$上的频率含量与 $\mathbf{X}$ 在 $f\in[0, 1/4]$上的频率含量有关.
- 另一方面，用小波滤波器$\{h_l\}$循环滤波 $\mathbf{X}$，然后用下 2 抽样得到序列 $\mathbf{W}_1=\mathbf{W}_{1,1}$，其在 $f\in[0, 1/2]$上的频率含量与 $\mathbf{X}$ 在 $f\in[1/4, 1/2]$上的频率含量有关，但是是逆序的(参见图 86).

这种对 $\mathbf{W}_{1,1}$表示的频率顺序与对 $\mathbf{X}$ 表示的频率顺序相比较的反转解释了反转高通和低通滤波操作，以得到在 $\mathbf{W}_{2,j}$ $(j=0, \cdots, 3)$中频率的合适对应的必要. 图 211 描述了量值平方离散傅里叶变换，如果小波和尺度滤波器是极好的高低通滤波器，那么量值平方离散傅里叶变换将会在图 210a 的流程图上的不同层里出现.

图 210b 表示出了看起来更直观的滤波操作的一个交替的顺序，因为在流程图里每次分裂的上部分总是进入 $G(\cdot)$，而下部分总是进入 $H(\cdot)$. 事实上，Wickerhauser(1994)认为这是自然顺序，但频率顺序没有直接反映到第二指标中.

上面的结果可以用一种明显的方式表现出来. 为了产生十分有用的连续顺序，我们必须遵循下面两个部分规则.

[1]如果 $\mathbf{W}_{j-1,n}$中的 n 是偶数，就用低通滤波器 $G(\cdot)$来得到 $\mathbf{W}_{j,2n}$，用高通滤波器$H(\cdot)$来得到 $\mathbf{W}_{j,2n+1}$. 用矩阵的记号，我们有

$$\mathbf{W}_{j,2n}=\mathcal{A}_j\mathbf{W}_{j-1,n} \quad 和 \quad \mathbf{W}_{j,2n+1}=\mathcal{B}_j\mathbf{W}_{j-1,n}. \tag{209a}$$

[2]如果 n 是奇数，就用 $H(\cdot)$得到 $\mathbf{W}_{j,2n}$，用 $G(\cdot)$得到 $\mathbf{W}_{j,2n+1}$：

$$\mathbf{W}_{j,2n}=\mathcal{B}_j\mathbf{W}_{j-1,n} \quad 和 \quad \mathbf{W}_{j,2n+1}=\mathcal{A}_j\mathbf{W}_{j-1,n}. \tag{209b}$$

用这个规则，我们可以在 j 层时构造 2^j 个向量，即 $\mathbf{W}_{j,n}$，$n=0, \cdots, 2^j-1$. 向量 $\mathbf{W}_{j,n}$是与区间$\mathcal{I}_{j,n}\equiv\left[\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}}\right]$上的标称频率相关的. 这个规则在图 212a 中阐述，一直分析到 3 层. 这个图是小波包表格(WP 表)或小波包树的一个例子. 因为我们分析的起点是时间序列 $\mathbf{X}$，所以很方便地定义 $\mathbf{W}_{0,0}\equiv\mathbf{X}$ 以使 $\mathbf{X}$ 与数对(j, n)相关，即$(0, 0)$.

取 $\mathbf{X}$ 到 $\mathbf{W}_{j,n}$的变换称为离散小波包变换(DWPT)，其中 $n=0, \cdots, 2^j-1$，$j\in\{0, \cdots, J\}$. 任何这样的变换都是规范正交的. 让我们看一下为何是这样的，

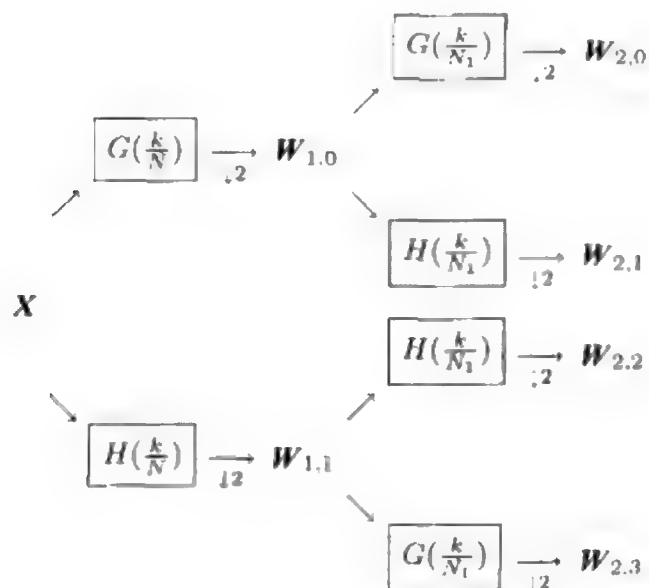


图 210a 描述 $\boldsymbol{X}$ 到 $\boldsymbol{W}_{2,0}$、$\boldsymbol{W}_{2,1}$、$\boldsymbol{W}_{2,2}$ 和 $\boldsymbol{W}_{2,3}$ 分析的流程图(连续顺序).

在上面选用了 $N_1 \equiv \frac{N}{2}$

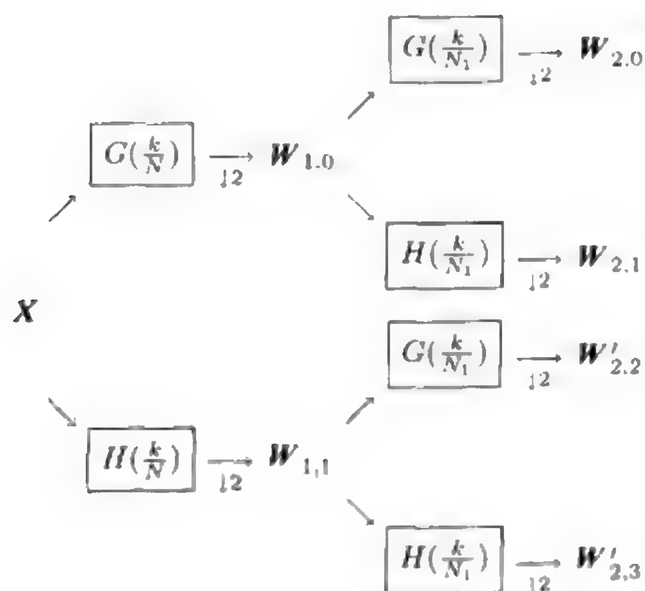


图 210b 描述 $\boldsymbol{X}$ 到 $\boldsymbol{W}_{2,0}$、$\boldsymbol{W}_{2,1}$、$\boldsymbol{W}'_{2,2}$ 和 $\boldsymbol{W}'_{2,3}$ 分析的流程图(自然顺序)

我们称 $\boldsymbol{W}_{j-1,n}$ 是“孩子”$\boldsymbol{W}_{j,2n}$ 和 $\boldsymbol{W}_{j,2n+1}$ 的“母体”. 从 $j-1$ 层到 j 层的离散小波包变换中，我们实质上将每一个母体用其两孩子代替. 如式(209a)和(209b)中所描述的，在母体结点的分裂是一个规范正交变换. 因为在 $j-1$ 层上的每个母体结点上这是正确的，故 $j-1$ 层的离散小波包变换的规范正交性隐含了在 j 层上变换

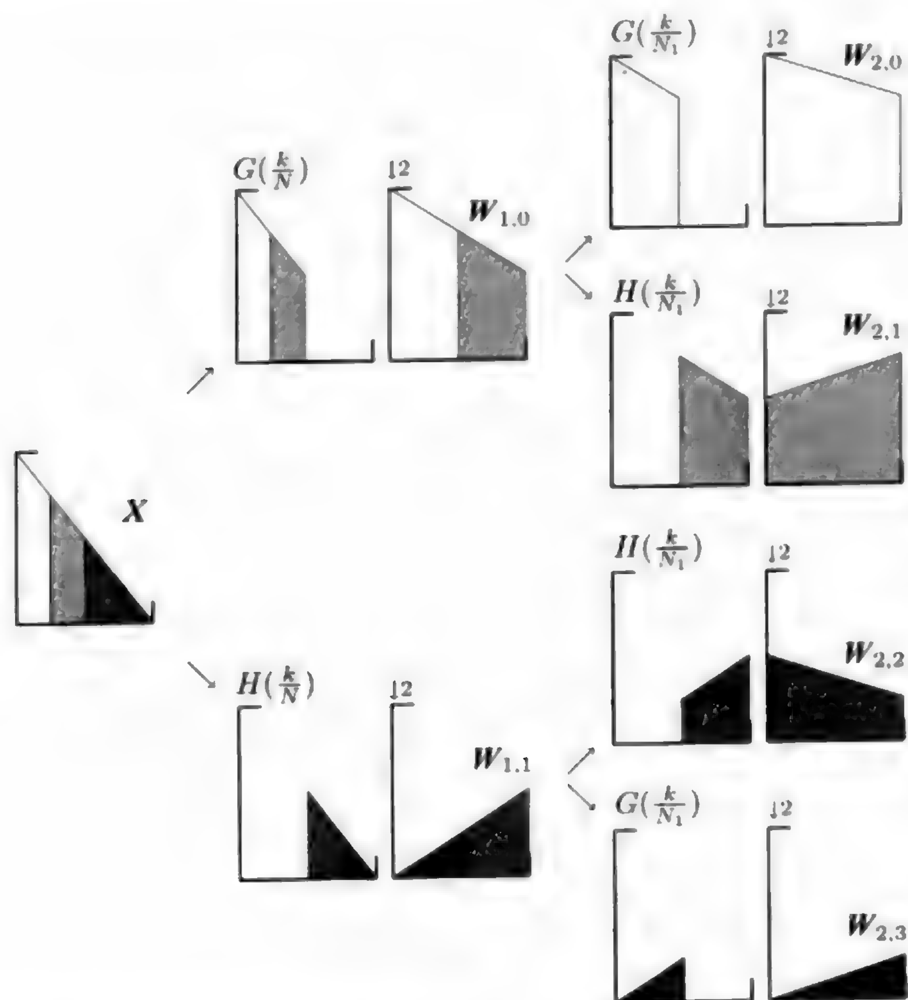


图 211　在频率域里用标准的低通和高通滤波器 $G(\cdot)$ 和 $H(\cdot)$ 滤波，然后又下抽样，得到的结果的描述(这个图与图 210a 的流程图是相同的). 长度为 N 的输入 $\boldsymbol{X}$ 的量的平方离散傅里叶变换，假定随着 f 从 0 移动到奈奎斯特频率而线性衰减. 用 $G(\cdot)$ 和 $H(\cdot)$ 滤波 $\boldsymbol{X}$ 产生一个半带序列，下抽样后变为全带序列($\boldsymbol{W}_{1,0}$ 或 $\boldsymbol{W}_{1,1}$，每个均有 $N/2$ 个点). 第二次滤波且下抽样产生 4 个新的全带序列，即 $\boldsymbol{W}_{2,n}$，$n=0$，1，2 和 3. 每个 $\boldsymbol{W}_{2,n}$ 长为 $N/4$ 且与 $\boldsymbol{X}$ 单一的四分之一带有关. 注意每次用 $H(\cdot)$ 滤波生成半带序列，其后的下抽样生成一个全带序列时，全带序列的频率内容与半带序列的频率含量反转；另一方面，用 $G(\cdot)$ 滤波不能产生这样的反转(参见 4.4 节里由 $H(\cdot)$ 引起的反转的解释)

的规范正交性.

形成结点指标的数对集合 (j, n) 由 $\mathcal{N} \equiv \{(j, n): j=0, \cdots, J_0;\ n=0, \cdots, 2^j-1\}$ 表示，其中我们选取任意 J_0 满足 $J_0 \leqslant J$(如果 $J_0 < J$，这对部分离散小波变换也是正确的，事实上我们可以放松约束 $N=2^J$ 并且允许 N 取 2^{J_0} 的任意整数倍). 对应着一个正交变换所形成的小波包系数指标的数对 (j, n) 将用 $\mathcal{C}$ 表示；例如，对给出 $\boldsymbol{W}_{3,0}$，…，$\boldsymbol{W}_{3,7}$ 的 $j=3$ 层的离散小波包变换，我们有 $\mathcal{C}=\{(3, n): n=0, \cdots, 7\}$. 很明显，$\mathcal{C} \subset \mathcal{N}$.

有意思的是，除了离散小波包变换，我们在小波包表格里能抽取许多其他的

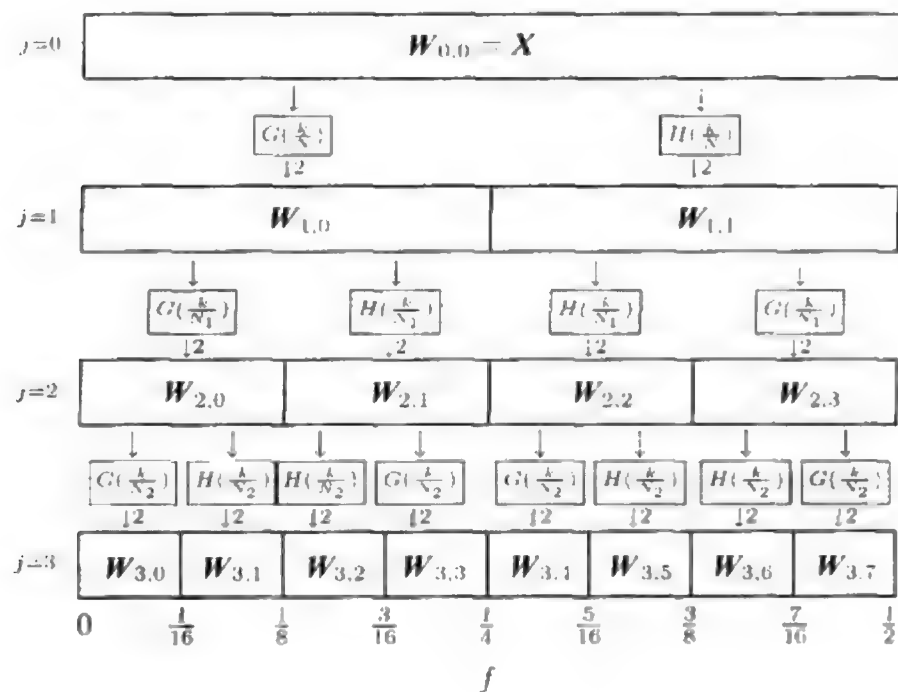


图 212a 描述 $\boldsymbol{X}$ 到 $\boldsymbol{W}_{3,0}$，…，$\boldsymbol{W}_{3,7}$ 分析的流程图（选用 $N_j = N/2^j$）

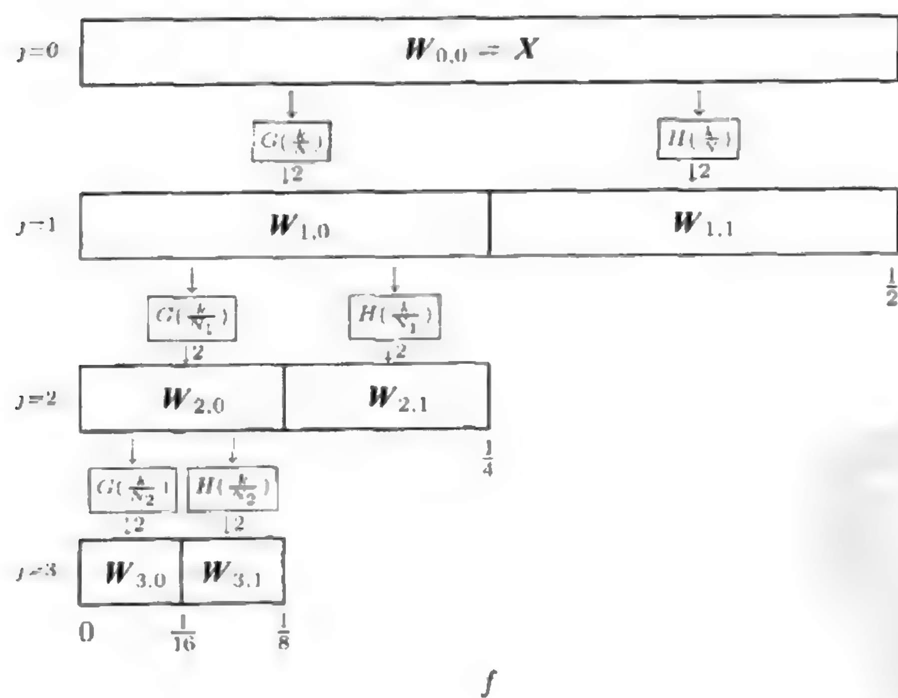


图 212b 描述 $\boldsymbol{X}$ 到 $\boldsymbol{W}_{3,0}$、$\boldsymbol{W}_{3,1}$、$\boldsymbol{W}_{2,1}$ 和 $\boldsymbol{W}_{1,1}$ 分析的流程图，这与 $J_0=3$ 层的部分离散小波变换是恒等的

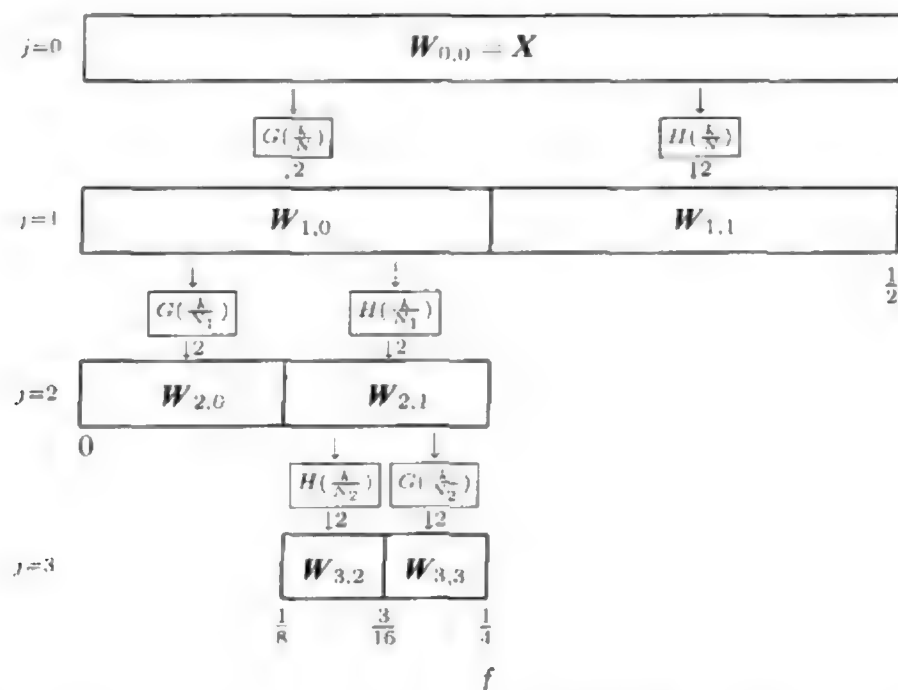


图 213　描述 $\boldsymbol{X}$ 到 $\boldsymbol{W}_{2,0}$、$\boldsymbol{W}_{3,2}$、$\boldsymbol{W}_{3,3}$ 和 $\boldsymbol{W}_{1,1}$ 分析的流程图，任意一个分开的二进分解

规范正交变换．例如，所有的部分离散小波变换可以从这样一个表格里拼合出来．这可以用图 212b 来描述，其中 $J_0=3$ 层部分离散小波变换展示出是由小波包表格里的 4 个结点组成的，即 $\mathcal{C}=\{(3,0),(3,1),(2,1),(1,1)\}$．注意这 4 个结点的标称频率区间形成 $\left[0,\frac{1}{2}\right]$ 的一个完全分划．更一般地，我们定义分开的二进分解，其在图 212a 的表格结构里的每个分裂点(可能的母体结点)上，我们或用 $G(\cdot)$ 和 $H(\cdot)$ 表示实行分裂，或根本就不分裂．这样的一个分解也是规范正交的，且也与 $\left[0,\frac{1}{2}\right]$ 的一个完全分划相关．图 213 表示出了任意的一个分开二进分解，这里 $\mathcal{C}$ 包含了数对(2，0)，(3，2)，(3，3)和(1，1)．

一旦我们选定了小波滤波器 $\{h_l\}$，那么小波包表格就可以完全特征化(相应的尺度滤波器 $\{g_l\}$ 就会由它与 $\{h_l\}$ 的正交镜像关系所确定，参见(75a))．任何从小波包表格里抽取的规范正交变换都可以由 $\mathcal{C}$ 里的数对特殊化．对图 213 的规范正交变换，我们有

$$\begin{bmatrix}\boldsymbol{W}_{1,1}\\ \boldsymbol{W}_{3,3}\\ \boldsymbol{W}_{3,2}\\ \boldsymbol{W}_{2,0}\end{bmatrix}=\begin{bmatrix}\mathcal{B}_1\\ \mathcal{A}_3\mathcal{B}_2\mathcal{A}_1\\ \mathcal{B}_3\mathcal{B}_2\mathcal{A}_1\\ \mathcal{A}_2\mathcal{A}_1\end{bmatrix}\boldsymbol{X}.$$

由于这个变换的规范正交性，我们有一个可加分解

$$\boldsymbol{X}=[\mathcal{B}_1^{\mathrm{T}},\mathcal{A}_1^{\mathrm{T}}\mathcal{B}_2^{\mathrm{T}}\mathcal{A}_3^{\mathrm{T}},\mathcal{A}_1^{\mathrm{T}}\mathcal{B}_2^{\mathrm{T}}\mathcal{B}_3^{\mathrm{T}},\mathcal{A}_1^{\mathrm{T}}\mathcal{A}_2^{\mathrm{T}}]\begin{bmatrix}\boldsymbol{W}_{1,1}\\ \boldsymbol{W}_{3,3}\\ \boldsymbol{W}_{3,2}\\ \boldsymbol{W}_{2,0}\end{bmatrix}$$

$$= \mathcal{B}_1^T W_{1,1} + \mathcal{A}_1^T \mathcal{B}_2^T \mathcal{A}_3^T W_{3,3} + \mathcal{A}_1^T \mathcal{B}_2^T \mathcal{B}_3^T W_{3,2} + \mathcal{A}_1^T \mathcal{A}_2^T W_{2,0}$$

和一个能量分解(方差分析)

$$\|X\|^2 = \|W_{1,1}\|^2 + \|W_{3,3}\|^2 + \|W_{3,2}\|^2 + \|W_{2,0}\|^2.$$

为了计算 $j=1, \cdots, J_0$ 层的离散小波包变换系数，我们在前一步里递推地滤波离散小波包变换系数. 令 $W_{j-1,n,t}$ 表示 $W_{j-1,n}$，$n=0, \cdots, 2^{j-1}-1$ 的第 t 个元素. 给定长度为 N_{j-1} 的这个向量，我们用两个部分规则来产生 $W_{j,n}$ 的元素，得到

$$W_{j,2n,t} \equiv \sum_{l=0}^{L-1} a_{n,l} W_{j-1,n,2t+1-l \bmod N_{j-1}},$$
$$t = 0, \cdots, N_j - 1,$$
$$W_{j,2n+1,t} \equiv \sum_{l=0}^{L-1} b_{n,l} W_{j-1,n,2t+1-l \bmod N_{j-1}},$$

其中

$$a_{n,l} = \begin{cases} g_l, & n \text{ 是偶数}; \\ h_l, & n \text{ 是奇数}, \end{cases} \qquad b_{n,l} = \begin{cases} h_l, & n \text{ 是偶数}; \\ g_l, & n \text{ 是奇数}. \end{cases}$$

练习[214] 证明上式一个等价的但是更简洁的写法是

$$W_{j,n,t} = \sum_{l=0}^{L-1} u_{n,l} W_{j-1,\lfloor \frac{n}{2} \rfloor, 2t+1-l \bmod N_{j-1}}, \quad t = 0, \cdots, N_j - 1, \tag{214a}$$

其中

$$u_{n,l} \equiv \begin{cases} g_l, & n \bmod 4 = 0 \text{ 或 } 3; \\ h_l, & n \bmod 4 = 1 \text{ 或 } 2, \end{cases} \tag{214b}$$

且 $\lfloor \cdot \rfloor$ 表示"整数部分"算子. ◁

我们也可以借助于滤波 X 适当下抽样来写 $W_{j,n}$(构造 X 到 $W_{j,n}$ 的滤波器的论证与 4.6 节离散小波变换的构造是十分类似的，这一部分的研究将说明怎样判断我们在这里的论断是否正确). 我们必须先在小波包表格里系统阐述对应于结点 (j, n) 的滤波器 $\{u_{j,n,l}: l=0, \cdots, L_j-1\}$ 的正确形式(像在式(96a)中，我们有 $L_j \equiv (2^j-1)(L-1)+1$). 假设我们令 $u_{1,0,l} \equiv g_l$ 和 $u_{1,1,l} \equiv h_l$，对 $j>1$ 的结点 (j, n)(与式(102)类似)，令

$$u_{j,n,l} \equiv \sum_{k=0}^{L-1} u_{n,k} u_{j-1,\lfloor \frac{n}{2} \rfloor, l-2^{j-1}k}, \quad l = 0, \cdots, L_j - 1 \tag{214c}$$

(通常，当 $l<0$ 或 $l \geq L_j$ 时，我们定义 $u_{j,n,l}=0$). 例如，对 $j=2$，

$$u_{2,0,l} = \sum_{k=0}^{L-1} u_{0,k} u_{1,0,l-2k} = \sum_{k=0}^{L-1} g_k g_{l-2k}$$

$$u_{2,1,l}=\sum_{k=0}^{L-1}u_{1,k}u_{1,0,l-2k}=\sum_{k=0}^{L-1}h_k g_{l-2k}$$

$$u_{2,2,l}=\sum_{k=0}^{L-1}u_{2,k}u_{1,1,l-2k}=\sum_{k=0}^{L-1}h_k h_{l-2k}$$

$$u_{2,3,l}=\sum_{k=0}^{L-1}u_{3,k}u_{1,1,l-2k}=\sum_{k=0}^{L-1}g_k h_{l-2k}.$$

那么对 $j=1,\cdots,J_0$，我们可以借助于滤波 $\boldsymbol{X}$ 并适当下抽样来写 $\boldsymbol{W}_{j,n}$ 的元素，即

$$W_{j,n,t}=\sum_{l=0}^{L_j-1}u_{j,n,l}X_{2^j[t+1]-1-l\bmod N},\quad t=0,1,\cdots,N_j-1. \tag{215a}$$

令 $U_{j,n}(\cdot)$表示$\{u_{j,n,l}\}$的传递函数，这个传递函数只依赖于尺度和小波滤波器的传递函数 $G(\cdot)$和 $H(\cdot)$. 为表示这种关系，与在 Coifman 等(1992)中一样，假设 $M_0(\cdot)\equiv G(\cdot)$和 $M_1(\cdot)\equiv H(\cdot)$. 也假设对小波包表格里的每个结点$(j,n)$，我们引入一个长度为 j，所有元素为 1 或 0 的向量 $\boldsymbol{c}_{j,n}$. 对 $j=1$，我们定义 $\boldsymbol{c}_{1,0}\equiv[0]$和 $\boldsymbol{c}_{1,1}\equiv[1]$. 对 $j>1$ 借助于 $\boldsymbol{c}_{j-1,n'}$，$n'=0,\cdots,2^{j-1}-1$，通过下面的两部分规则循环定义 $\boldsymbol{c}_{j,n}$，$n=0,\cdots,2^j-1$.

[1]如果 $n\bmod 4=0$ 或 3，我们通过在 $\boldsymbol{c}_{j-1,\lfloor\frac{n}{2}\rfloor}$的末尾添加一个 0 产生 $\boldsymbol{c}_{j,n}$.

[2]如果 $n\bmod 4=1$ 或 2，我们在 $\boldsymbol{c}_{j-1,\lfloor\frac{n}{2}\rfloor}$的末尾添加一个 1 产生 $\boldsymbol{c}_{j,n}$.

看一下怎样用这个规则，例如，假设我们想构造 $j=4$，$n=7$ 的 $\boldsymbol{c}_{j,n}$. 因为 $7\bmod 4=3$，故这个法则的部分[1]表明给 $\boldsymbol{c}_{3,3}$ 添加一个 0 来构造 $\boldsymbol{c}_{4,7}$，为构造 $\boldsymbol{c}_{3,3}$，注意 $3\bmod 4=3$，故部分[1]的另一个应用表明给 $\boldsymbol{c}_{2,1}$ 添加 0，最后因为 $1\bmod 4=1$，故用法则的部分[2]来构造 $\boldsymbol{c}_{2,1}$，即给 $\boldsymbol{c}_{1,0}\equiv[0]$添加 1，综合这些结果，我们有 $\boldsymbol{c}_{4,7}\equiv[0,1,0,0]^T$.

设 $c_{j,n,m}$表示 $\boldsymbol{c}_{j,n}$的第 m 个元素，则

$$U_{j,n}(f)=\prod_{m=0}^{j-1}M_{c_{j,n,m}}(2^m f). \tag{215b}$$

向量 $\boldsymbol{c}_{j,n}$可以用图 216 中的排列次序的小波包表格来解释. 作为例子，考虑与 $\boldsymbol{c}_{3,3}=[c_{3,3,0},c_{3,3,1},c_{3,3,2}]^T=[0,1,0]^T$ 有关的 $U_{3,3}(f)$. 然后我们有

$$U_{3,3}(f)=M_0(f)M_1(2f)M_0(4f)=G(f)H(2f)G(4f).$$

对 $j=3$，$n=0,\cdots,7$ 的 LA(8)滤波器，平方增益函数$|U_{j,n}(\cdot)|^2$ 在图 217 中被表示出. 标称的"理想"的带通是由垂线标明的. 例如，对 $U_{3,3}(\cdot)$的标称带通是$\left[\frac{3}{16},\frac{1}{4}\right]$.

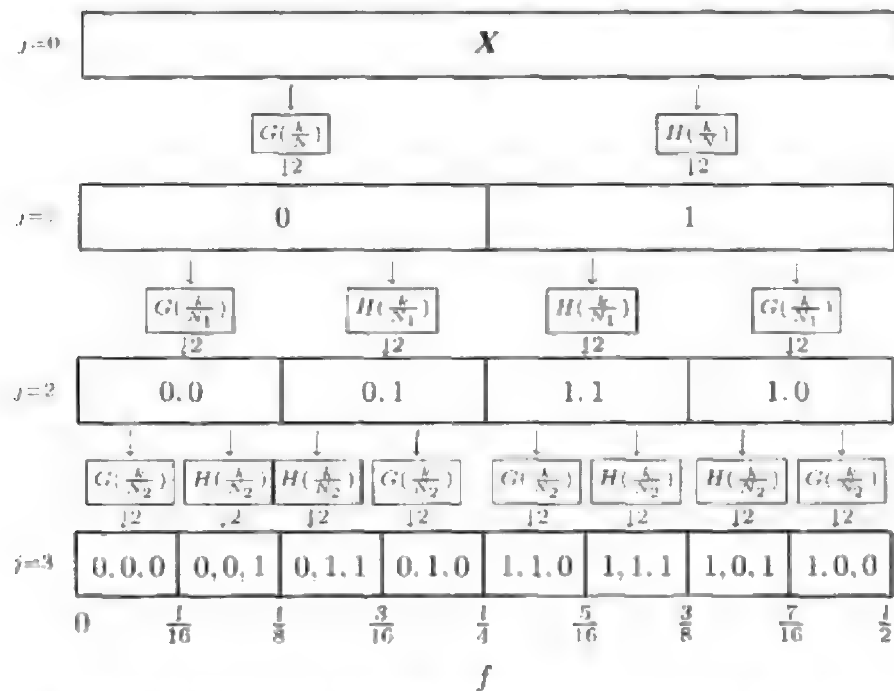


图 216 对 $j=1$，2 和 3 层的排列次序的小波包表，向量 $c_{j,n}$ 的解释(0 或 1 分别表示尺度(低通)或小波(高通)滤波器的使用)．注意如果在 $j-1$ 层的母结点到 j 层的子结点进行时，我们要求使用 $G(\cdot)$，那么就添加一个 0 到母的 $c_{j-1,\lfloor\frac{n}{2}\rfloor}$ 以得到子的 $c_{j,n}$；另一方面，如果我们用 $H(\cdot)$，那么就添加一个 1．还注意到，当从左边到右边穿过行 $j=2$ 或 3 移动且选出每个 $c_{j,n}$ 的最后一个元素时，我们得到了模型"0，1，1，0"或这个模型的继续重复．(对通常的 $j\geqslant 2$，搜集 $c_{j,n}$，$n=0$，…，2^j-1 的最后元素将会产生 2^{j-2} 个"0，1，1，0"的重复)

最后，回忆在第 j 层离散小波包变换的每个 $\boldsymbol{W}_{j,n}$ 标称是与带宽 $\frac{1}{2^{j+1}}$ 有关的，即与之相关的频率区间 $\mathcal{I}_{j,n}$ 的宽度．也注意到 $\boldsymbol{W}_{j,0}$ 是与 j 层部分离散小波变换的尺度系数的向量 $\boldsymbol{V}_j$ 完全相同的．这些尺度系数是与尺度 $\lambda_j=2^j$ 相关的(这个量可以用序列的自相关宽度的记号来论证，参见式(103)和围绕它的讨论)．因此，我们可以将 λ_j 的标称的"时宽"与 $\boldsymbol{W}_{j,0}$ 联系起来．因为用来在 j 层创造所有 $\boldsymbol{W}_{j,n}$ 的等价滤波有完全相同的宽度 L_j 和相同的标称带宽，故我们可以论证地使相同的时宽 λ_j 隶属于每个 $\boldsymbol{W}_{j,n}$(注意，因为在每个 $\boldsymbol{W}_{j,n}$ 里，有 $N_j=\frac{N}{2^j}=\frac{N}{\lambda_j}$ 个系数，每个系数的 λ_j 的标称宽度可以产生总的时间范围 $N_j\times\lambda_j=N$；即 $\boldsymbol{W}_{j,n}$ 的系数集体覆盖了 $\boldsymbol{X}$ 提供的整个时间，按照直观是合理的)．对 $j=0$，时宽是 1，然而带宽是 $\frac{1}{2}$，这在给定 $j=0$ 层的离散小波包变换是一个恒等变换时，是很容易被发现的．如果我们考虑另一个极端的例子，即长为 $N=2^J$，时宽为 N，而带宽为 $\frac{1}{2N}$ 的序列的 J 层的离散小波包变换．由于没有任何一个等价滤波器 $\{u_{j,n,l}\}$ 在任何意义下是局部化的，故时宽是合理的．对所有的 j 层，时宽和带宽之积是一个常数 $\frac{1}{2}$，并且和在时间域和频率域之间的许多"互反性关系"保持一致(例如，参见 Bracewell，1978，第 8 章或 Percival and Walden，1993，第 3 章)．

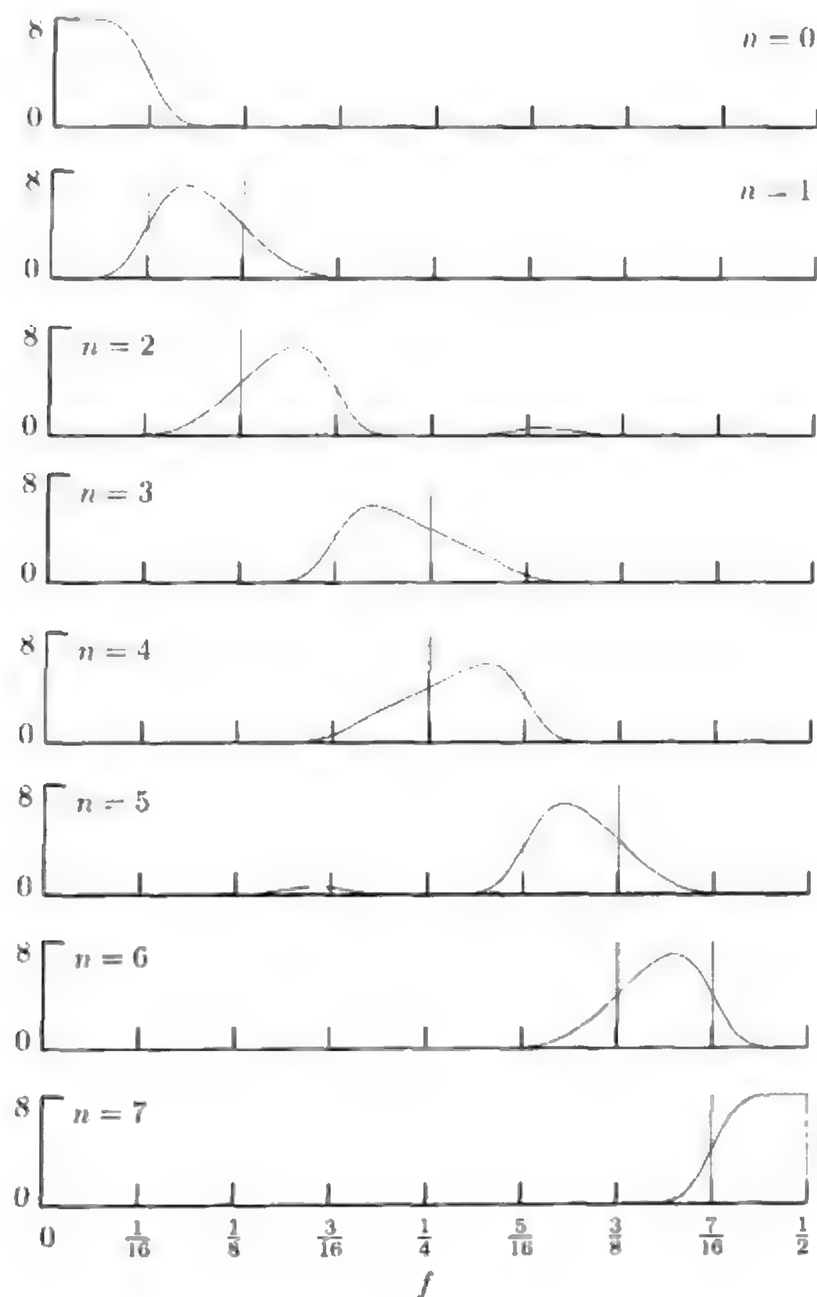


图 217　基于 LA(8) 小波和尺度滤波器，对于 $j=3$ 和 $n=0, \cdots, 7$ 的平方增益函数 $|U_{j,n}(\cdot)|^2$. 标称的"理想"通带是由垂线标记的

6.1 节的评论与扩展

[1]上面给出的用来构造二进值向量 $\boldsymbol{c}_{j,n}$ 的两部分规则，允许我们决定哪个序列能引出带有平方增益函数的，在频率区间 $\mathcal{I}_{j,n} \equiv \left[\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}}\right]$ 上是标称带宽的滤波器，这个序列必须是具有形式 $G(2^m f)$ 和/或 $H(2^m f)$ 的传递函数的滤波器卷积得到的. 像规则陈述的 $\mathcal{I}_{j,n}$ 与 $\boldsymbol{c}_{j,n}$ 相连接的一个更为正式的证明，参见 Walden and Contreras Cristan(1998a).

[2]假设我们有长度为 $N=2^J$ 的时间序列，并且使用哈尔小波构造 J 层的离散小波包变换. 这个变换事实上恒等于文献中著名的(规范正交)沃尔什变换. 这种变换对 $N=16$ 的情况的列向量在图 218 表示. 注意像离散傅里叶变换(参见图 47)，而不像 $j<J$ 层的离散小波包变换，对于沃尔什变换的基向量在时间上

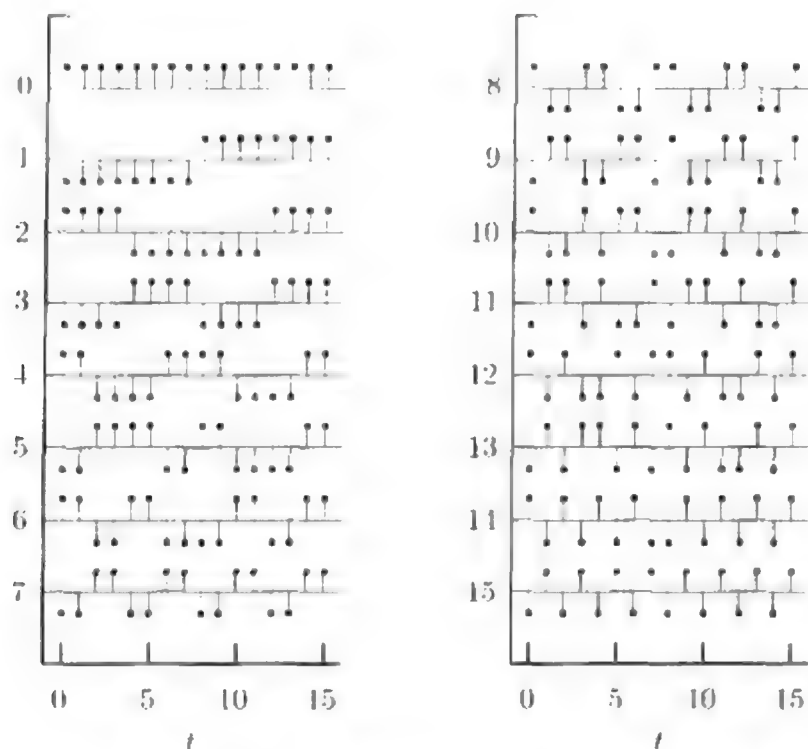


图 218 使用哈尔小波滤波器构造的 $J=4$ 层的 $N=16$ 离散小波包变换的行向量，这个变换与著名沃尔什(Walsh)变换是一样的. 基向量按顺时针次序被展示，以使第 n 行(的转置)和时间序列 $\mathbf{X}$ 的内积产生系数向量 $\mathbf{W}_{4,n}$ 里的单个元素

没有被局部化. 沃尔什变换作为离散傅里叶变换的一个对手/补充，在一些细节上被研究过了. 更仔细地，参见 Morettin(1981)，Beauchamp(1984)或 Stoffer(1991)；用沃尔什变换的有意思的应用，参见 Stoffer 等(1988)和 Lanning and Johnson(1983).

[3]如我们注意到的，第 j 层的离散小波包变换产生能被划分成 2^j 个向量的变换系数，即 $\mathbf{W}_{j,n}$，$n=0,\cdots,2^j-1$，第 n 个这样的向量是与区间 $\mathcal{I}_{j,n}=\left[\frac{n}{2^{j+1}},\frac{n+1}{2^{j+1}}\right]$ 上的标称频率相关的. 因为每个 $\mathcal{I}_{j,n}$ 的宽度都是相同的，故离散小波包变换能有效地将频率范围 $\left[0,\frac{1}{2}\right]$ 划分为 2^j 个整齐的隔开区间. 由于这种划分和离散小波包变换是以一种与离散小波变换十分相似的方式计算的，一些研究者把离散小波包变换称作是“规则的离散小波变换”(例如，参见 Medley 等，1994).

6.2 例子：太阳物理数据的离散小波包变换

在过程和现象的一个复杂重叠太阳系的太阳球面上，磁场的物理学是由发生在太阳系的和在不同长尺度的太阳风里，与观察点附近局部更多的一样的点构成的. 尤里西斯(Ulysses)宇宙太空计划，第一时间里，提供了太阳系磁极地区的观察. 这些地区是在太阳活动循环的一个最小相位期间内被测量的，且因此被从磁极日冕洞出来的快速太阳风流动支配. 数据覆盖的无先例高比例(超过 95%)——随着尤里西斯的太阳系磁极轨道——提供了太阳球面磁场的独特高质量记录.

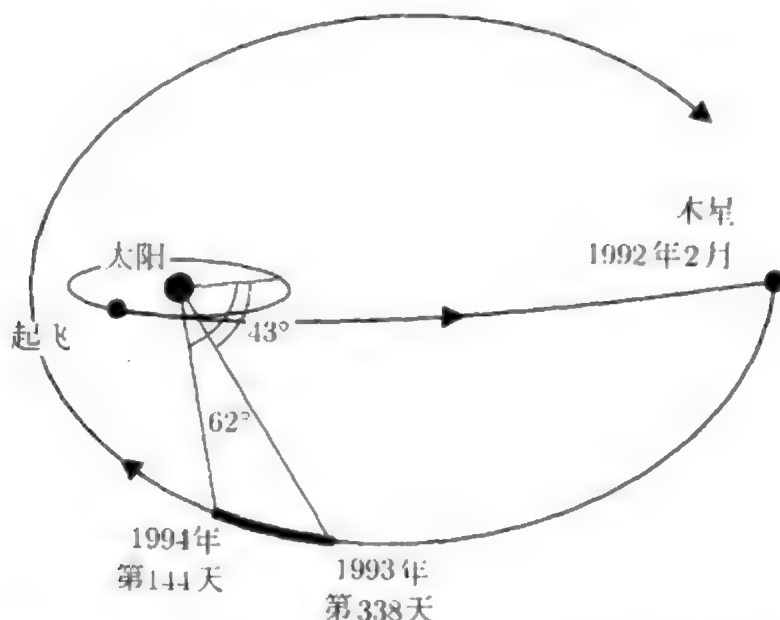


图 219　尤里西斯宇宙飞船的轨道．宇宙飞船起飞后，在地球活动的平面上运行，直到木星附近．木星的引力场将之投掷到太阳附近的磁极轨道上．图 222 底部绘出的那部分数据是在宇宙飞船上从日照计纬度南43°到太阳以南的南62°时搜集的——在改变纬度时，飞船也缩短了与太阳的距离．这里收集到感兴趣的数据后，飞船继续它的太阳磁极轨道，但现在是在太阳的北面运行(经伦敦 Queen Mary 和 Westfield 学院的 T. Horbury 的说明许可)

作为分析的有用数据是由磁场大小的每小时 $N=4\ 096$ 个平均值的自由分割间隔组成的，这个磁场大小是从 1993 年(12 月 4 日)的第 338 天宇宙时间 21 时到 1994 年(3 月 24 日)的第 144 天宇宙时间 12 时记录的，且包括了尤里西斯正在记录从太阳系南极的发射物这段时间(参见图 219)．这个序列在图 222 底部被绘出(这个相同图的一个扩张版本在图 235 的底部被表示出来)．为方便，我们假定时间变量是在白昼时被测出的，因此抽样区间是 $\Delta t=1/24$ 天(故频率在每天的循环中被测量且范围从 0 上升到每天 12 周的奈奎斯特频率之内变化)．测量的单位是毫微特斯拉(nanoteslas)，即10^{-9}特斯拉，用 nT 表示．和图 219 中表述的一样，在这个时间序列被记录的区间内，尤里西斯从大约南43°到南62°改变了日照计的纬度，并且它的日心范围从大约 4 到 3 个天文单位变化(由定义，一个天文单位与地球到太阳的平均距离是相等的)．这个向南的半球面数据显示出一些大的尺度结构．我们注意到的 4 个震荡波在图 222 里被标以 a(1994 年 2 月 12 日)、b(1994 年 2 月 26 日)、c(1994 年 3 月 10 日)、d(1994 年 4 月 3 日)，我们将测验这些高纬度特点的时间/频率特性．

行星际的震动波被典型地分为“上旋转”震动或“变换”震动(例如，González-Esparza 等，1996)．上旋转震动是由快速太阳风(发始于太阳的日冕洞)在行星介质里超过慢速太阳风的相互作用所产生的．多次发生的相互作用区域被称为上旋转相互作用域(CIR)，并且再现每个太阳旋转．变换震动被认为是由快速日冕的物质喷射(CME)所产生的．分类方法的细节在 González-Esparza 等(1996)给出．

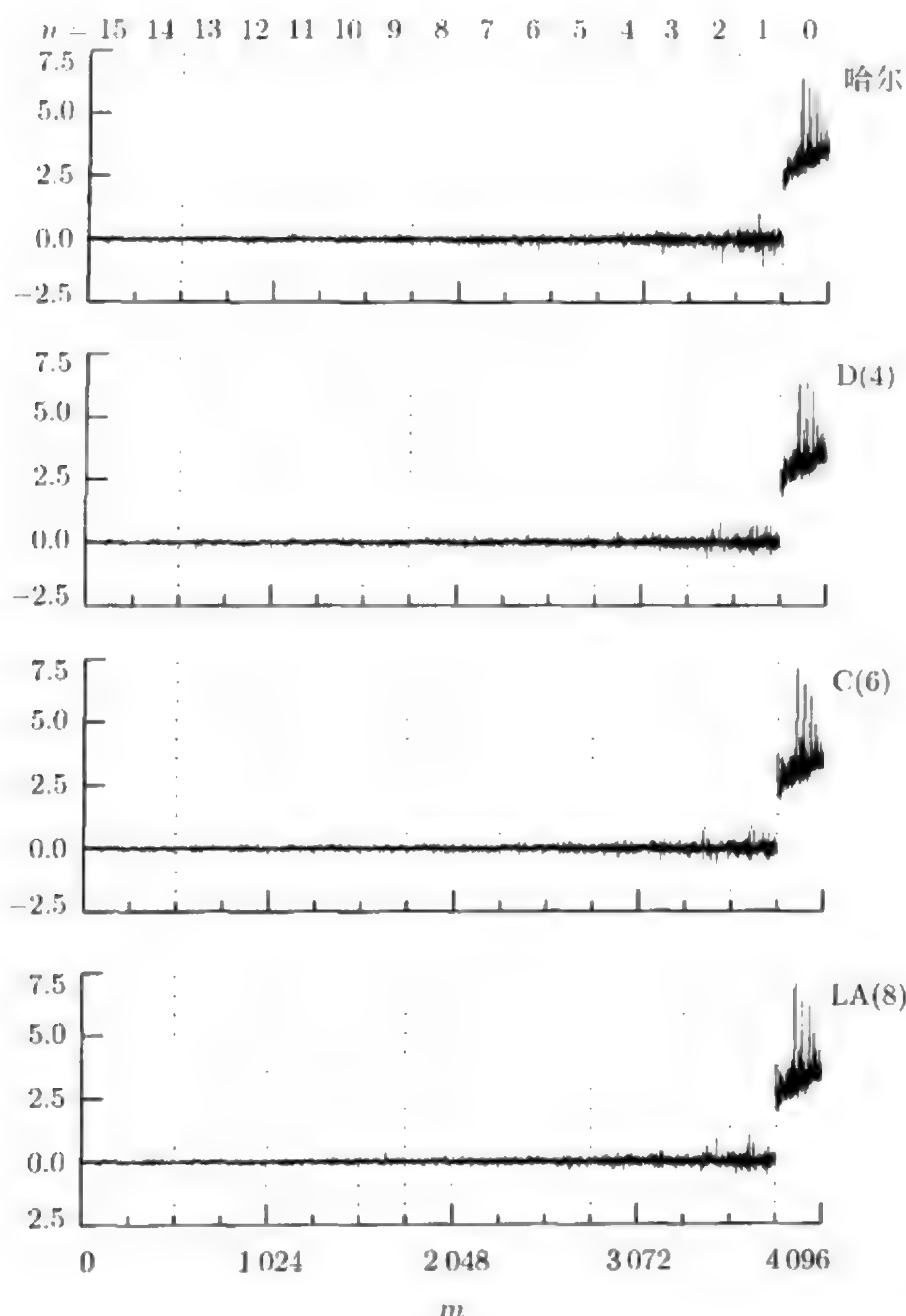


图 220　使用哈尔、D(4)、C(6)和 LA(8)小波对太阳物理时间序列的 $j=4$ 层的离散小波包变换系数 $\boldsymbol{W}$. $\boldsymbol{W}$ 的元素 W_m 相对 $m=0, \cdots, N-1=4\ 095$ 被绘出图. 15 条垂直的虚线描绘出了 $\boldsymbol{W}$ 的 16 个子向量，即从左到右，$\boldsymbol{W}_{4,n}$，$n=15, 14, \cdots, 0$

Balogh 等(1995)在图 222 中分类了与日冕物质喷射相关的作为变换的前两个震动波和后两个反复出现的上旋转相互作用域. 然而，Zurbuchen 等(1996)把第一个、第三个和第四个看作是反复出现的.

由于这个序列有着随时间变化的分量，故它是说明离散小波包变换分析性能的一个好的选择. 图 220 表示出了基于哈尔、D(4)、C(6)和 LA(8)小波的 $j=4$ 层的离散小波包变换系数向量 $\boldsymbol{W}$ 的元素 W_m. 垂直的虚线表示了每个 $\boldsymbol{W}$ 到子向量 $\boldsymbol{W}_{4,n}$，$n=0, \cdots, 2^j-1=15$ 的分划(注意 $\boldsymbol{W}_{4,15}$ 是图中左上方的，而 $\boldsymbol{W}_{4,0}$ 是右上方的). 系

数向量 $\mathbf{W}$ 和时间序列 $\mathbf{X}$ 有相同的元素个数. 我们有 $\|\mathbf{W}\|^2=\|\mathbf{X}\|^2$ 是因为离散小波包变换是规范正交变换的. 每个子向量 $\mathbf{W}_{4,n}$ 有相同的元素个数，即 $N_j=\frac{N}{2^j}=256$. 子向量 $\mathbf{W}_{4,n}$ 在实际频率的区间 $\frac{1}{\Delta t}\mathcal{I}_{4,n}=[n/(32\Delta t),(n+1)/(32\Delta t)]$ 上跟踪时间上的变化. 图表明了，系数的大小通常是随着减少的频率而增加，而最大系数属于 $\mathbf{W}_{4,0}$(与频率的最低区间有关). 注意当 $\mathbf{W}_{4,n}(n>0)$ 中的系数在零附近波动时，那些在 $\mathbf{W}_{4,0}$ 的系数就在正值处波动. 这可以通过练习[6.6]的结果来解释，它告诉我们，$\mathbf{W}_{4,0}$ 的抽样平均值与 $4\bar{X}$ 相等，其中，和通常一样，$\bar{X}$ 是 $\mathbf{X}$ 的抽样平均值(如图 222 所表示的，这里显然是正的).

在给定的 $\mathbf{W}_{4,n}$，$n\geqslant3$ 内，图 220 里对于所有 4 个小波包变换通过时间的变差出现满意的相似和性质一样，但在与三个最低频率带相关的子向量中，尤其在 $\mathbf{W}_{4,0}$ 中，有一些显著的脉冲. 我们探索图 222 上部里不同带的脉冲之间的时间关系，其中 LA(8)离散小波包变换子向量的每个元素相对考虑到 LA(8)滤波器的近似零相位特性的一个时间而被绘出来. 这些调节依赖于 6.5 节讨论的平移 $\nu_{4,n}$，它被用来连接 $\mathbf{W}_{4,n}$ 的第 t 个元素 $W_{4,n,t}$ 和时间 $t_0+(2^4(t+1)-1-|\nu_{4,n}|\bmod N)\Delta t$，其中 t_0 是与 X_0 有关的实际时间(这个公式与 4.8 节对离散小波变换得出的是完全同样的). 产生次序(按照这个次序其元素实际被绘制成图)的 $\mathbf{W}_{4,n}$ 的循环平移在图 222 的右端表示——例如，$\mathcal{T}^{-2}\mathbf{W}_{4,0}$ 表明元素 $W_{4,0,1}$ 和 $W_{4,0,2}$ 是分别与最早的和最晚的时间相关的. 每个子向量的开头和结尾处的粗垂线描绘了边界系数的轮廓：在这两条线以外被绘出的系数在一定程度上被离散小波包变换假设的循环性所影响. 我们可以用 4.11 节描述的和离散小波变换十分相同的方式确定 $\mathbf{W}_{4,n}$ 的哪一个元素是边界系数(对每个 $\mathbf{W}_{4,n}$，其边界系数的个数是由 L_4' 给出，表 136 说明 LA(8)滤波器的边界系数个数为 6). 图上面部分的虚垂线表明了 4 个由 Balogh 等(1995)确定的震动波结构的位置——这些定时是通过从原始记录里在(不是每小时的平均值)数据中直观上寻找恰当的特性所决定的. 对于用 LA(8)小波使之可能的相位定位，我们能在这些震动波的时间周围看到几个频率带内大系数的配置；然而我们将在 6.7 节看到，在离散小波包变换里固有的下抽样损害我们的能力以辨别这些主带的特色.

6.3 最好基算法

像在 6.1 节我们注意到的，我们可以从小波包表格中提取直到某一 J_0 层的许多不同的规范正交变换. 每个可能的规范正交变换都是一个分开的二进分解，由定义是频率区间 $\left[0,\frac{1}{2}\right]$ 的一个分划(这些变换包括——但不局限——部分离散小波变换和 $j=1,\cdots,J_0$ 层的离散小波包变换). 我们怎样能确定这许多变换中的哪一个在某种意义上对某个特定的时间序列 $\mathbf{X}$ 是最优的呢? 等价地，给定小波滤波器的选择，哪个 $\mathcal{C}\subset\mathcal{N}$ 在某种意义上是最优的?

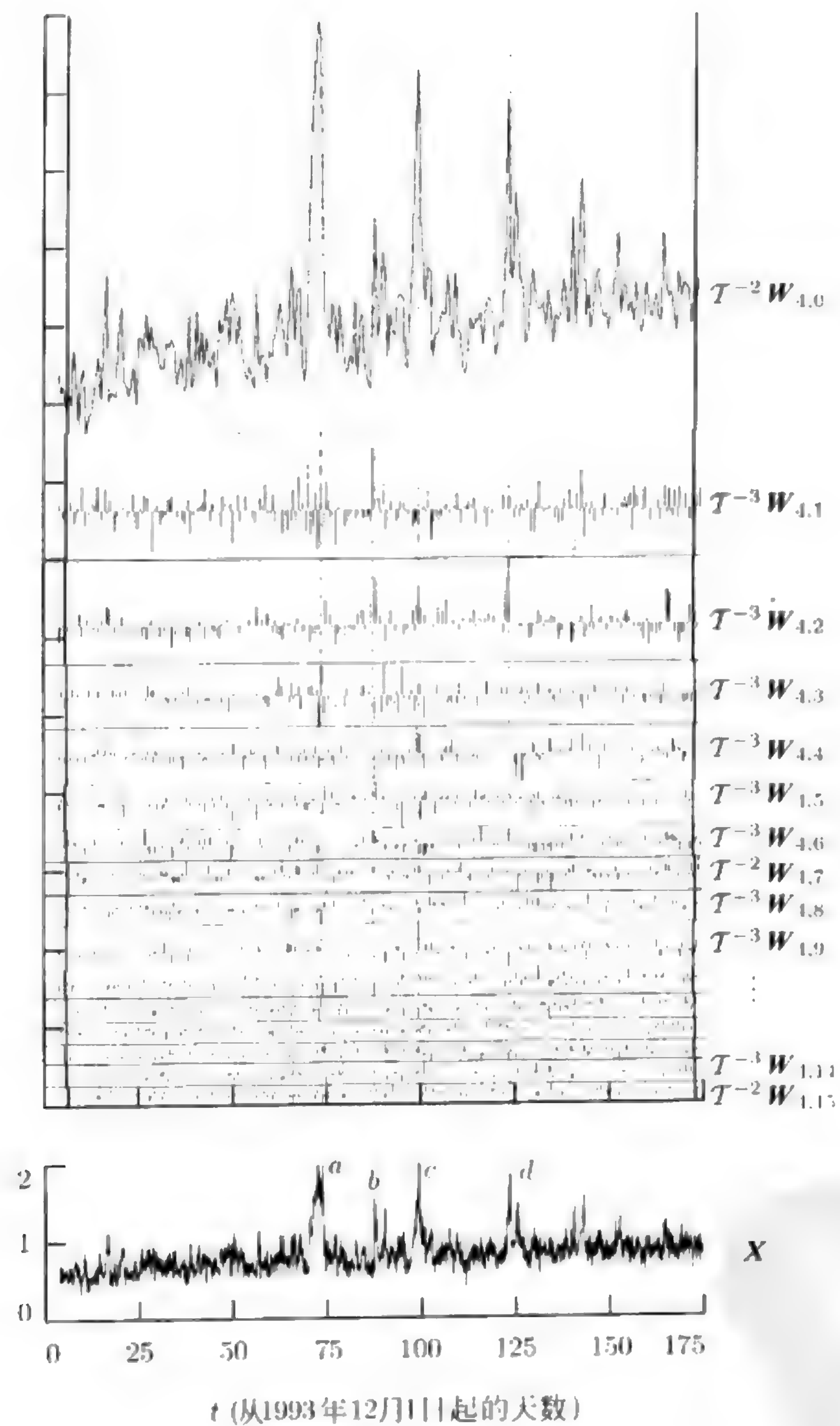


图 222　太阳物理时间序列的 $j=4$ 层的 LA(8)离散小波包变换系数(详情参见正文)

Coifman and Wickerhauser(1992)在“最好基算法”的命名下解决了这个问题，最好基算法包括下面两个基本步骤.

[1]给定一个 J_0 层的小波包表格，然后对每个$(j, n)\in\mathcal{N}$，赋给 $\mathbf{W}_{j,n}$ 一个价值 $M(\mathbf{W}_{j,n})$，其中 $M(\cdot)$是一个形式为

$$M(\mathbf{W}_{j,n}) \equiv \sum_{t=0}^{N_j-1} m(|W_{j,n,t}|)$$

的相加的价值泛函，且 $m(\cdot)$是定义在$[0, \infty)$有 $m(0)=0$ 的实值函数.

[2]从小波包表格中提取的“最优”正交变换是

$$\min_{\mathcal{C}} \sum_{(j,n)\in\mathcal{C}} M(\mathbf{W}_{j,n})$$

的解，即我们寻求由$\mathcal{C}\subset\mathcal{N}$特征化的正交变换以使所有数对$(j, n)\in\mathcal{C}$的价值和最小.

许多相加的价值泛函已经提出(例如，参见 Wickerhauser，1994，第 8 章). 这里有三个例子.

[1]$-\ell^2\log(\ell^2)$范数，也称为“熵”信息价值泛函，其中

$$m(|\overline{W}_{j,n,t}|) = \begin{cases} -\overline{W}^2_{j,n,t}\log(\overline{W}^2_{j,n,t}), & W_{j,n,t}\neq 0; \\ 0, & W_{j,n,t}=0, \end{cases}$$

且 $\overline{W}_{j,n,t}\equiv\dfrac{W_{j,n,t}}{\|\mathbf{X}\|}$. 这种量与序列的熵之间有一种单调关系.

[2]阈值泛函，即超出特定阈值 δ 的项的个数，即定义

$$m(|W_{j,n,t}|) = \begin{cases} 1, & |W_{j,n,t}|>\delta; \\ 0, & \text{其他}. \end{cases}$$

[3]ℓ_p 信息价值泛函，对于 $0<p<2$，定义为 $m(|W_{j,n,t}|)=|W_{j,n,t}|^p$. 如果使用这个价值泛函，那么 $M^{\frac{1}{p}}(\mathbf{W}_{j,n})$就给出了序列的 ℓ_p 范数.

练习[223] 解释一下为何价值泛函 $m(|W_{j,n,t}|)=|W_{j,n,t}|^2$ 对于选取一个“最优”正交变换没有什么用处. ◁

为得到这些价值泛函是一些什么概念，让我们来考虑下面的简单例子. 假设用哈尔小波由长度为 $N=8$ 的时间序列 $\mathbf{X}$ 形成层为 $j=1, 2, 3$ 的离散小波包变换的系数：

$$\begin{bmatrix} \mathbf{W}_{1,1} \\ \mathbf{W}_{1,0} \end{bmatrix} = \begin{bmatrix} \mathcal{B}_1 \\ \mathcal{A}_1 \end{bmatrix}\mathbf{X}, \quad \begin{bmatrix} \mathbf{W}_{2,3} \\ \mathbf{W}_{2,2} \\ \mathbf{W}_{2,1} \\ \mathbf{W}_{2,0} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_2\mathcal{B}_1 \\ \mathcal{B}_2\mathcal{B}_1 \\ \mathcal{B}_2\mathcal{A}_1 \\ \mathcal{A}_2\mathcal{A}_1 \end{bmatrix}\mathbf{X} \quad \text{和} \quad \begin{bmatrix} \mathbf{W}_{3,7} \\ \mathbf{W}_{3,6} \\ \mathbf{W}_{3,5} \\ \mathbf{W}_{3,4} \\ \mathbf{W}_{3,3} \\ \mathbf{W}_{3,2} \\ \mathbf{W}_{3,1} \\ \mathbf{W}_{3,0} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_3\mathcal{A}_2\mathcal{B}_1 \\ \mathcal{B}_3\mathcal{A}_2\mathcal{B}_1 \\ \mathcal{B}_3\mathcal{B}_2\mathcal{B}_1 \\ \mathcal{A}_3\mathcal{B}_2\mathcal{B}_1 \\ \mathcal{A}_3\mathcal{B}_2\mathcal{A}_1 \\ \mathcal{B}_3\mathcal{B}_2\mathcal{A}_1 \\ \mathcal{B}_3\mathcal{A}_2\mathcal{A}_1 \\ \mathcal{A}_3\mathcal{A}_2\mathcal{A}_1 \end{bmatrix}\mathbf{X}.$$

<table>
<tr><td>$j=0$</td><td colspan="8">$X^T=[2,0,-1,1,0,0,-2,2]$</td></tr>
<tr><td>$j=1$</td><td colspan="4">$W_{1,0}^T=[\sqrt{2},0,0,0]$</td><td colspan="4">$W_{1,1}^T=[-\sqrt{2},\sqrt{2},0,\sqrt{8}]$</td></tr>
<tr><td>$j=2$</td><td colspan="2">$W_{2,0}^T=[1,0]$</td><td colspan="2">$W_{2,1}^T=[-1,0]$</td><td colspan="2">$W_{2,2}^T=[2,2]$</td><td colspan="2">$W_{2,3}^T=[0,2]$</td></tr>
<tr><td>$j=3$</td><td>$[\frac{1}{\sqrt{2}}]$</td><td>$[-\frac{1}{\sqrt{2}}]$</td><td>$[\frac{1}{\sqrt{2}}]$</td><td>$[-\frac{1}{\sqrt{2}}]$</td><td>$[\sqrt{8}]$</td><td>$[0]$</td><td>$[\sqrt{2}]$</td><td>$[\sqrt{2}]$</td></tr>
</table>

$0 \quad \frac{1}{16} \quad \frac{1}{8} \quad \frac{3}{16} \quad \frac{1}{4} \quad \frac{5}{16} \quad \frac{3}{8} \quad \frac{7}{16} \quad \frac{1}{2}$

f

图 224 对长为 $N=8$ 的时间序列 $\boldsymbol{X}$，层为 $j=1$、2 和 3 的哈尔离散小波包变换系数 $\boldsymbol{W}_{j,n}$. 序列 $\boldsymbol{X}$ 是由三个基向量的线形组合构造的，每个基向量来自于 3 个层中的一个(参见式(224)). 在上面，与形成 $\boldsymbol{X}$ 时用到的向量相对应离散小波包变换系数已被下划线了(在 $\boldsymbol{W}_{1,0}$、$\boldsymbol{W}_{2,3}$ 和 $\boldsymbol{W}_{3,4}$ 中各一个)

其中 $\mathcal{A}_j$ 和 $\mathcal{B}_j$ 维数为 $\frac{8}{2^j}\times\frac{8}{2^{j-1}}$. 我们从 $\mathcal{A}_1$ 的第一行向量，$\mathcal{A}_2\mathcal{B}_1$ 的第二行向量和 $\mathcal{A}_3\mathcal{B}_2\mathcal{B}_1$ 的行向量(其转置)的线性组合来构造 $\boldsymbol{X}$：

$$\boldsymbol{X}=\sqrt{2}\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\\0\\0\\0\\0\\0\end{bmatrix}+2\begin{bmatrix}0\\0\\0\\0\\-1/2\\1/2\\-1/2\\1/2\end{bmatrix}+\sqrt{8}\begin{bmatrix}1/\sqrt{8}\\-1/\sqrt{8}\\-1/\sqrt{8}\\1/\sqrt{8}\\1/\sqrt{8}\\-1/\sqrt{8}\\-1/\sqrt{8}\\1/\sqrt{8}\end{bmatrix}=\begin{bmatrix}2\\0\\-1\\1\\0\\0\\-2\\2\end{bmatrix}. \tag{224}$$

通过观察可以看到，用来构造 $\boldsymbol{X}$ 的 3 个向量是相互正交的，但每个都由来自不同阶的离散小波包变换. 图 224 给出了对直到阶 $J_0=3$ 的基于哈尔的离散小波包变换，其 $\boldsymbol{W}_{j,n}$ 的内容——这里我们强调用来构造 $\boldsymbol{X}$ 的 3 个向量所形成的系数(每个都在 $\boldsymbol{W}_{1,0}$、$\boldsymbol{W}_{2,3}$ 和 $\boldsymbol{W}_{3,4}$).

对于构造 $\boldsymbol{X}$ 的一个"最好基"的一个直观的想法是确信包含着我们形成 $\boldsymbol{X}$ 时的三个规范正交向量——如果这些包含进去，就能用三个非零系数简洁地表示出 $\boldsymbol{X}$. 考虑系数向量 $\boldsymbol{W}_{1,0}$，它的第一个元素是用式(224)表示的第一个向量形成的. 因为 $\boldsymbol{W}_{1,0}$ 对应着频率区间 $\left[0,\ \frac{1}{4}\right]$，在形成分开的二进分解时，它的竞争者或 $\boldsymbol{W}_{2,0}$ 与 $\boldsymbol{W}_{2,1}$ 一起，或 $\boldsymbol{W}_{3,0}$ 与 $\boldsymbol{W}_{3,1}$、$\boldsymbol{W}_{3,2}$ 和 $\boldsymbol{W}_{3,3}$ 一起. 因此，一个合理的相加的价值泛函应该分配给 $\left[\sqrt{2},\ 0,\ 0,\ 0\right]$ 或分配给 $[1,\ 0]$ 和 $[-1,\ 0]$(即 $\boldsymbol{W}_{2,0}$ 与 $\boldsymbol{W}_{2,1}$ 一起组成的组) 或分配给 $\left[\frac{1}{\sqrt{2}}\right]$、$\left[-\frac{1}{\sqrt{2}}\right]$、$\left[\frac{1}{\sqrt{2}}\right]$ 和 $\left[-\frac{1}{\sqrt{2}}\right]$(即 $j=3$ 的系数参与竞争的一组)使有更低的价值的一组. 从品质上说，从其他两组竞争系数中区别 $\left[\sqrt{2},\ 0,\ 0,\ 0\right]$ 的特性是因为其系数中的一个在数量上比其他的更突出；换句话说，它的系数是非齐次的.

$j=0$	1.45							
$j=1$	0.28				0.88			
$j=2$	0.19		0.19		0.72		0.36	
$j=3$	**<u>0.12</u>**	**<u>0.12</u>**	**<u>0.12</u>**	**<u>0.12</u>**	**<u>0.32</u>**	**<u>0.00</u>**	**<u>0.28</u>**	**<u>0.28</u>**

f: 0, $\frac{1}{16}$, $\frac{1}{8}$, $\frac{3}{16}$, $\frac{1}{4}$, $\frac{5}{16}$, $\frac{3}{8}$, $\frac{7}{16}$, $\frac{1}{2}$

图 225 对图 224 表示出小波包表格的价表. 上面的表是用$-\ell^2\log(\ell^2)$范数作为一个价值泛函被计算的$M(\mathbf{W}_{j,n})$的值. 最好基算法的初始步指定我们应该以某种方式在低行标记出那些值——这里我们用下划线和粗体字标记

接着我们看一下这个三个价值泛函怎样处理齐次系数的一个简单例子，譬如，$\mathbf{W}_h=\left[\pm\frac{\varepsilon}{2},\ \pm\frac{\varepsilon}{2},\ \pm\frac{\varepsilon}{2},\ \pm\frac{\varepsilon}{2}\right]^{\mathrm{T}}$ 相对于非齐次系数相反的一个向量，譬如，$\mathbf{W}_i=[\pm\varepsilon,\ 0,\ 0,\ 0]^{\mathrm{T}}$(注意这两个向量的能量是相同的，这对任何一个给定频率区间$\mathcal{I}_{j,n}$的处于竞争中的任何系数集都是正确的).

[1]使用$-\ell^2\log(\ell^2)$范数(为方便，假设 $\mathbf{W}_h$ 和 $\mathbf{W}_i$ 是从$\|\mathbf{X}\|=1$ 的序列得出的)，我们有

$$M(\mathbf{W}_h)=-4\cdot\frac{\varepsilon^2}{4}\log\left(\frac{\varepsilon^2}{4}\right)=-\varepsilon^2(\log(\varepsilon^2)-\log(4))>-\varepsilon^2\log(\varepsilon^2)=M(\mathbf{W}_i),$$

因此，如所期望的，$\mathbf{W}_i$ 被分配了比 $\mathbf{W}_h$ 更低的价值.

[2]如果我们使用，譬如 $\delta=\frac{\varepsilon}{4}$，作为阈值泛函的阈值，那么 $M(\mathbf{W}_h)=4$，然而$M(\mathbf{W}_i)=1$，故 $\mathbf{W}_i$ 再次有一个比较低的价值；另一方面，如果我们置，譬如$\delta=\frac{3\varepsilon}{4}$，那么$M(\mathbf{W}_h)=0$ 而 $M(\mathbf{W}_i)=1$，这与我们想要的相反. 因此，阈值泛函能够由于选择合适 δ 的问题可被巧妙地运用.

[3]如果我们使用 ℓ_p 信息价值泛函，那么，对任意的 $0<p<2$，有

$$M(\mathbf{W}_h)=4\left|\frac{\varepsilon}{2}\right|^p>|\varepsilon|^p=M(\mathbf{W}_i),$$

又一次与我们想要的是一致的.

因此，假设我们能在阈值泛函里适当地设置 δ，所有三个价值泛函分配给非齐次系数的向量一个较低的价值.

一旦价值泛函被选定，就可以给小波包表格里的每个向量分配一个价值. 这样一个“价值表格”(基于$-\ell^2\log(\ell^2)$范数)的例子在图 225 里对应着图 224 的小波包表格中给出. 最好基算法(Coifman and Wickerhauser，1992；Wickerhauser，1994)以一种简单的方法进行，像图 225 和 226 里举例说明的.

[1]我们以某种方式在表格底部标记所有结点的价值(图 225 里用下划线且用黑体打印). 我们从检查结点的低行开始.

[2]我们将每对孩子结点的价值和与它们双亲结点的价值进行比较，然后做下列中的一个：

$j=0$	0.96							
$j=1$	**0.28**				0.68			
$j=2$	**0.19**		**0.19**		0.32		**0.36**	
$j=3$	**0.12**	**0.12**	**0.12**	**0.12**	**0.32**	**0.00**	**0.28**	**0.28**

$0 \quad \frac{1}{16} \quad \frac{1}{8} \quad \frac{3}{16} \quad \frac{1}{4} \quad \frac{5}{16} \quad \frac{3}{8} \quad \frac{7}{16} \quad \frac{1}{2}$

f

图 226　最好基算法的最后一步. 最好基变换是用阴影箱指定的. 由于直观上是合理的，故选择的变换包括了用来形成 $\boldsymbol{X}$ 的三个基向量

(a)如果双亲结点的价值比孩子结点的价值之和低，就给双亲结点划记号(即在图中的价值上标上下划线且用黑体字打印).

(b)如果孩子结点的价值之和比双亲结点的价值低，就用孩子结点的价值之和代替双亲结的价值(什么新记号也不做).

[3]顺着表格向上，我们对每一层重复步骤[2]，最后的结果在图 226 中给出.

[4]一旦我们到达了表格的顶端，就可以沿着表格向下看那些被标记过的结点. 对应着一个分离的二进分解的被标记过的最上边的结点定义了最好基变换.

对我们的简单例子，最好基变换在图 226 里用阴影方格表示出来并且是由对应于离散小波包变换的系数 $\boldsymbol{W}_{1,0}$、$\boldsymbol{W}_{3,4}$、$\boldsymbol{W}_{3,5}$ 和 $\boldsymbol{W}_{2,3}$ 的基向量组成. 这一选择在直觉上是合理的，因为 $\boldsymbol{X}$ 是用由 $\boldsymbol{W}_{1,0}$、$\boldsymbol{W}_{3,4}$ 和 $\boldsymbol{W}_{2,3}$ 的基向量形成的(这三组基向量不能形成分离的二进分解，但这样的一个分解的形成是基于包括添加形成的向量 $\boldsymbol{W}_{3,5}$).

6.3 节的评论与扩展

[1]最好基算法是一个“底部提升”算法，因此需要直到 J_0 层的整个小波包表格的计算. Taswell(1996)讨论了一种能要求考虑相当少的计算的“顶部下降”方法. 虽然这种方法选择了一个在理论上次最优的基，但计算机实验表明用各种标准测量的次最优和最好基之间，几乎都没有实际差别.

6.4　例子：太阳物理数据的最好基

这里我们将最好基算法应用到太阳磁场的量值数据上(这些在 6.2 节描述，并且在图 222 中图示). 因为这个序列的抽样区间是 $\Delta t=\frac{1}{24}$天，故我们用$\frac{1}{\Delta t}=24$乘以两个端点，将标准化的频率区间 $\mathcal{I}_{j,n}=\left[\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}}\right]$变换到有实际意义的频率区间$\left[\frac{n}{2^{j+1}\Delta t}, \frac{n+1}{2^{j+1}\Delta t}\right]$，其中这个区间里的任何频率均以每天周测量的. 如果

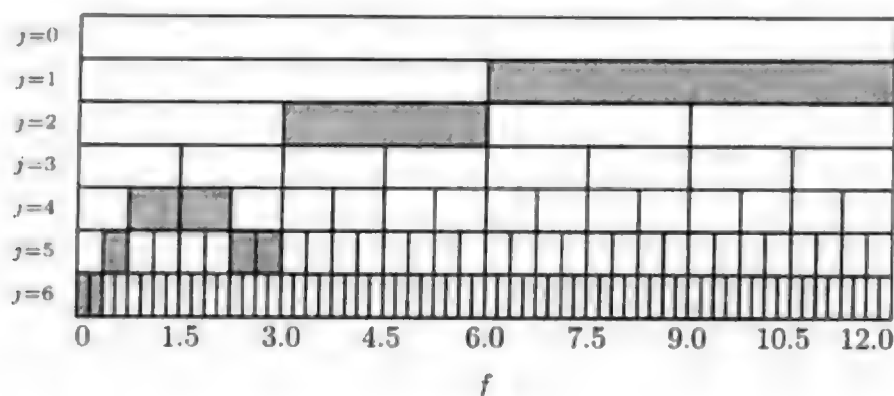


图 227 对图 222 的太阳磁场量值数据的最好基变换. 这个变换是基于 LA(8)小波滤波器和基于 $-\ell^2\log(\ell^2)$范数的价值泛函

用 LA(8)小波滤波器构造这个序列的至到 $J_0=6$ 层的小波包表格，然后如果使用基于 $-\ell^2\log(\ell^2)$范数的一个价值泛函应用最好基算法，我们就会得到图 227 描述的最好基变换. 变换是由 $j=1$ 层离散小波包变换的 $\frac{N}{2}$ 个向量、$j=2$ 层的 $\frac{N}{4}$ 个向量、$j=4$ 层的 $\frac{N}{8}$ 个向量、$j=5$ 层的 $\frac{3N}{32}$ 个向量和 $j=6$ 层的 $\frac{N}{32}$ 个向量构成的. 注意用一种与离散小波变换有些类似的方式，这个变换一般是随着频率的降低而增加它的频率分辨率——频率区间[2.25, 3.0]周/天是这个一般模式的一个例外(我们将在 6.7 节重新返回到这个变换，连同为这个序列的极大重叠离散小波变换分析选择一个层以寻找作为时间函数跟踪频率变化).

让我们现在考虑最好基算法的两种可能适合太阳磁极量值数据的改善. 第一，从图 222 底部的图中可以看到，在这个时间序列里有一个明显的向上的漂移，这是有问题的，因为与离散小波变换一样，离散小波包变换也利用了循环滤波. 在时间序列的开始和末端之间的非匹配能够导致几个十分大的边界系数. 因为最好基算法运用了试图选择非齐次系数集的度量标准，这些反常的系数是十分令人苦恼的. 为了缓和这样的忧虑，让我们使用反射边界条件重新计算这个时间序列的最好基变换；即与式(140)描述的一样，我们将序列扩展到它原来长度的两倍并且分析这个增加过的序列. 选择的基在图 228 的顶端表示. 如果将这个基与图 227 里用周期边界条件选择的基进行比较，就会看到，直到 $j=3$ 层的基向量都是相同的，除了有一些明显的差别以外. 边界条件的选择因此能够影响最好基算法.

第二个改善试图缓解离散小波包变换对于我们将时间序列“分组”的依赖(这个问题与在 5.1 节里我们注意过的离散小波变换的问题是类似的). 这个依赖性影响着最好基算法，如果我们循环平移一个时间序列，对这个平移过的序列，其最好基变换没有必要与原来序列的变换相同. 这种考虑已经使一些研究人员提出了“平移不变性”的一种形式，它能通过使用与在最好基算法里用到的相同的价值

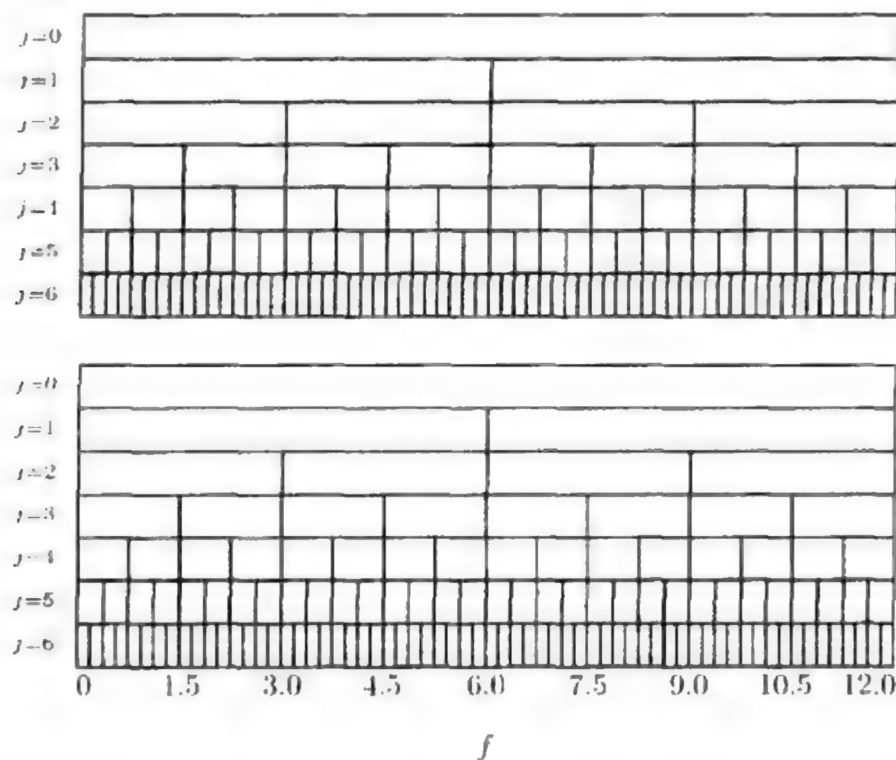


图 228　对于太阳磁极量值数据，基于反映边界条件的最好基变换（顶部）和对相同数据又一次使用反射边界条件的最好平移基变换（底部）．这两个变换都是基于图 227 里用到的相同小波滤波器和价值泛函．最好平移基算法选择了平移 $m=52$（即 $\mathcal{T}^{52}$）作为最好平移

泛函在不同循环平移中来有效地选择（例如，参见 Del Marco and Weiss，1997 以及其中的参考文献）．这就导致了最好基算法的下述修改，我们称之为“最好平移基算法”．我们首先形成对 $\boldsymbol{X}$ 的直到 J_0 层的小波包表格，然后应用最好基算法，并且观察与选择的基有关的总价值（作为图 226 表示的例子，这等于 4 个阴影箱里的数目之和）．下一步，我们对每个循环平移序列 $\mathcal{T}^m\boldsymbol{X}$（$m=1$，…，$2^{J_0}-1$）重复上面的操作（注意我们仅需要考虑这些平移：任何 $m<0$ 或 $m\geqslant 2^{J_0}$ 将产生一个小波包表格，它的分量 $\boldsymbol{W}_{j,n}$ 将仅仅是对 $m \bmod 2^{J_0}$ 的平移成分的循环平移变型——因此对 $\mathcal{T}^m\boldsymbol{X}$ 和 $\mathcal{T}^{m \bmod 2^{J_0}}\boldsymbol{X}$ 的得到的价值表格是相同的，并且因此导致了相同的最好基变换）．最好平移基变换定义为与产生最小总价值的平移相关的最好基变换．

当应用到太阳磁场量值数据时，最好平移基变换是与平移 $m=52$ 相关的，且在图 228 的底部展示出（和这个图的顶部一样，我们已经使用了 LA(8) 小波，基于 $-\ell^2\log(\ell^2)$ 范数的一个价值泛函和对称边界条件）．注意选择的变换的频率分辨率随着频率的减少而增大（一个类似的模式几乎对图 227 和图 228 顶部所描述的变换均保持）．对 $J_0=6$，有总数为 64 的平移（包括 $m=0$）能潜在地引起不同的最好基变换．对正在研究的时间序列，所有 64 个最好基变换是十分类似的，至多在图中描述的有 3 个较小的变动．

6.5 小波包滤波器的时间平移

在 4.8 节中，我们观察了 Daubechies 的最接近对称小波和尺度滤波器的相位特性，并且在 4.9 节里对 coiflet 滤波器做出了相同的事情. 特别地，我们得到了作为零相位滤波器的近似输出（下抽样）而给出离散小波变换系数图的时间平移（当在第 5 章做出极大重叠离散小波变换系数的图时，我们再一次用到了这些时间平移）. 这里我们将这种方法延伸到用来创造离散小波包变换的滤波器上（也是这个变换的一种"极大重叠"形式——极大重叠离散小波包变换——将在 6.6 节讨论）.

得到恰当时间平移的关键是给出滤波器 $\{u_{j,n,l}\}$ 的传递函数 $U_{j,n}(\cdot)$ 的式(215b). 在定义 $U_{j,n}(\cdot)$的乘积中的每一个因子能写为

$$M_{c_{j,n,m}}(2^m f) = |M_{c_{j,n,m}}(2^m f)| e^{i\theta_{c_{j,n,m}}(2^m f)},$$

其中

$$\theta_{c_{j,n,m}}(f) = \begin{cases} \theta^{(G)}(f), & c_{j,n,m} = 0; \\ \theta^{(H)}(f), & c_{j,n,m} = 1. \end{cases}$$

从式(112d)

$$\theta^{(G)}(f) \approx 2\pi f \nu \quad \text{和} \quad \theta^{(H)}(f) \approx -2\pi f(L-1+\nu),$$

这里，由式(112e)和 4.9 节的讨论，

$$\nu = \begin{cases} -\frac{L}{2}+1, & \text{对带有偶数}\frac{L}{2}\text{的最接近对称滤波器；} \\ -\frac{L}{2}, & \text{对 LA(10) 和 LA(18) 滤波器；} \\ -\frac{L}{2}+2, & \text{对 LA(14) 滤波器；} \\ -\frac{2L}{3}+1, & \text{对 coiflet 滤波器}(L=6,12,\cdots,30). \end{cases} \tag{229a}$$

注意 ν 总是负的（我们提醒自己有这样一个用 $-|\nu|$ 代替 ν 的事实）. 因此相对应 $U_{j,n}(\cdot)$的相位函数可被写为

$$\sum_{m=0}^{j-1}\theta_{c_{j,n,m}}(2^m f) \approx 2\pi f[\nu S_{j,n,0} - (L-1+\nu)S_{j,n,1}],$$

其中

$$S_{j,n,0} \equiv \sum_{m=0}^{j-1}(1-c_{j,n,m})2^m \quad \text{和} \quad S_{j,n,1} \equiv \sum_{m=0}^{j-1} c_{j,n,m}2^m.$$

如果我们注意到总有

$$S_{j,n,0} + S_{j,n,1} = \sum_{m=0}^{j-1} 2^m = 2^j - 1, \tag{229b}$$

那么我们可写为

$$\sum_{m=0}^{j-1}\theta_{c_{j,n,m}}(2^m f) \approx 2\pi f[\nu(2^j-1-S_{j,n,1}) - (L-1+\nu)S_{j,n,1}] \equiv 2\pi f \nu_{j,n},$$

其中 $\nu_{j,n} \equiv \nu(2^j-1) - S_{j,n,1}(2\nu+L-1)$.

表 230

j	n	$c_{j,n}$	$S_{j,n,0}$	$S_{j,n,1}$	$\|\nu_{j,n}\|$	LA(8)	$\|p_{j,n}\|$	LA(8)
1	0	0	1	0	$\frac{L_1}{2}-1$	3	$e\{g_l\}$	2.8
	1	1	0	1	$\frac{L_1}{2}+0$	4	$e\{h_l\}$	4.2
2	0	0，0	3	0	$\frac{L_2}{2}-2$	9	$3e\{g_l\}$	8.5
	1	0，1	1	2	$\frac{L_2}{2}+0$	11	$e\{g_l\}+2e\{h_l\}$	11.2
	2	1，1	0	3	$\frac{L_2}{2}+1$	12	$3e\{h_l\}$	12.5
	3	1，0	2	1	$\frac{L_2}{2}-1$	10	$2e\{g_l\}+e\{h_l\}$	9.8
3	0	0，0，0	7	0	$\frac{L_3}{2}-4$	21	$7e\{g_l\}$	19.9
	1	0，0，1	3	4	$\frac{L_3}{2}+0$	25	$3e\{g_l\}+4e\{h_l\}$	25.2
	2	0，1，1	1	6	$\frac{L_3}{2}+2$	27	$e\{g_l\}+6e\{h_l\}$	27.8
	3	0，1，0	5	2	$\frac{L_3}{2}-2$	23	$5e\{g_l\}+2e\{h_l\}$	22.5
	4	1，1，0	4	3	$\frac{L_3}{2}-1$	24	$4e\{g_l\}+3e\{h_l\}$	23.8
	5	1，1，1	0	7	$\frac{L_3}{2}+3$	28	$7e\{h_l\}$	29.1
	6	1，0，1	2	5	$\frac{L_3}{2}+1$	26	$2e\{g_l\}+5e\{h_l\}$	26.5
	7	1，0，0	6	1	$\frac{L_3}{2}-3$	22	$6e\{g_l\}+e\{h_l\}$	21.2

注：需要前移$|\nu_{j,n}|$的形式，基于偶数$\frac{L}{2}$的最接近对称滤波器，用滤波器$\{u_{j,n,l}\}$来取得近似零相位输出．这里令$j=1$，2和3且$n=0$，…，2^j-1．第7列给出了LA(8)滤波器的特殊值(参见6.5节评论和扩展中对后两列的讨论)．

练习[230] 证明对我们考虑过的所有最接近对称和coiflet滤波器，都有$\nu_{j,n}<0$. ◁

因此，我们通过用$|\nu_{j,n}|=-\nu_{j,n}$单位提升$\{u_{j,n,l}\}$以得到一个零相位滤波器的最接近的逼近．

现在滤波器$\{u_{j,n,l}\}$有宽度$L_j\equiv(2^j-1)(L-1)+1$，因此我们能写

$$\nu_{j,n}=\left(\frac{L_j-1}{L-1}\right)\nu-S_{j,n,1}[2\nu+L-1].$$

用上式和式(229a)中给出的对各种滤波器中ν的表示式，我们得到(经过一些代数运算后)：

$$\nu_{j,n}=\begin{cases}-\frac{L_j}{2}+(2^{j-1}-S_{j,n,1}), & \text{对偶数}\frac{L}{2}\text{的最接近对称滤波器；}\\ -\frac{L_j}{2}-(2^{j-1}-S_{j,n,1})+1, & \text{对 LA(10) 和 LA(18)；}\\ -\frac{L_j}{2}+3(2^{j-1}-S_{j,n,1})-1, & \text{对 LA(14)；}\\ -\frac{L_j}{2}-\frac{L-3}{3}\left(2^{j-1}-S_{j,n,1}-\frac{1}{2}\right)+\frac{1}{2}, & \text{对 coiflet 滤波器.}\end{cases}\tag{230}$$

上面的式子对列表格是很方便的，因为它是根据滤波器半长度的偏差来表示$\nu_{j,n}$的．对于偶数$\frac{L}{2}$和对于$j=1$，…，3而$n=0$，…，2^j-1的最接近对称滤波器的情况，要求使最接近零相位的滤波器$\{u_{j,n,l}\}$的提升$|\nu_{j,n}|$的代数形式在表230的第

六列中给出．第七列给出了对 LA(8)滤波器的 $|\nu_{j,n}|$ 实际的数字值(最后两列在下边的评论和扩展中讨论)．

6.5 节的评论与扩展

[1]像在 4.8 节的评论和扩展[2]中所述的一样，Wickerhauser(1994，第 171—341 页)和 Hess Nielsen and Wickerhauser(1996)使用一个滤波器"中心能量"的记号系统地阐述了适用于大范围的滤波器设计的提升(这些设计包括了 Daubechies 的极值相位和最接近对称滤波器及 coiflet 滤波器)．这些想法还能使用到离散小波包变换滤波器 $\{u_{j,n,l}\}$ 上．恰当的前移 $|p_{j,n}|$ 是由

$$|p_{j,n}| = S_{j,n,0}e\{g_l\} + S_{j,n,1}e\{h_l\}$$

给出的，其中，因为它们都有单位能量，$\{g_l\}$ 和 $\{h_l\}$ 的能量中心分别是由

$$e\{g_l\} = \sum_{l=0}^{L-1} lg_l \quad 和 \quad e\{h_l\} = \sum_{l=0}^{L-1} lh_l$$

给出的．对带有偶数 $\frac{L}{2}$ 的最接近对称滤波器，其 $|p_{j,n}|$ 的形式在表 230 的倒数第二列中表示，且对 LA(8)滤波器特定的值在最后一列中表出(对这个滤波器，计算表明 $e\{g_l\} \approx 2.8464$ 和 $e\{h_l\} \approx 4.1536$)．注意 $|p_{j,n}|$ 四舍五入到最近的整数后，与 $|\nu_{j,n}|$ 至多在一致性上不同，因此确定前移的两种方法是很不相似的．

6.6 极大重叠离散小波包变换

在第 5 章中，我们探讨了作为离散小波变换的一种选择的极大重叠离散小波变换(MODWT)．极大重叠离散小波变换有好几个离散小波变换所不具有的可取特性，包括对一个时间序列的假设起点的较小灵敏度，任何抽样尺寸 N 的适用性和与零相位滤波器相关的多分辨分析(在转换到极大重叠离散小波变换时，我们付出的主要代价是一个额外的计算负担和正交性的丢失)．可以将极大重叠观点直接延伸到定义一个极大重叠离散小波包变换(MODWPT)中．这里，我们概略这个变换的主要观点和结果(正式的证明简化或省略了，因为他们是由第 5 章已经提出的极大重叠离散小波变换的接近的翻版)．

我们根据离散小波包变换的副本通过定义极大重叠离散小波包变换滤波器 $\{\tilde{u}_{n,l}\}$ 和 $\{\tilde{u}_{j,n,l}\}$ 开始，即

$$\tilde{u}_{n,l} \equiv \frac{u_{n,l}}{\sqrt{2}} \quad 和 \quad \tilde{u}_{j,n,l} \equiv \frac{u_{j,n,l}}{2^{j/2}},$$

其中 $u_{n,l}$ 和 $u_{j,n,l}$ 在式(214b)和(214c)中给出．用这些滤波器，我们可以用两种等价方法中的一种定义极大重叠离散小波包变换的 j 层系数 $\widetilde{W}_{j,n,t}$．第一种需要用 $\{\tilde{u}_{j,n,l}\}$ 直接滤波 $\mathbf{X}$：

$$\widetilde{W}_{j,n,t} \equiv \sum_{l=0}^{L_j-1} \tilde{u}_{j,n,l} X_{t-l \bmod N}, \quad t = 0,1,\cdots,N-1. \tag{231a}$$

第二种是一种递推方案，假设我们已经有 $j-1$ 层系数：

$$\widetilde{W}_{j,n,t} \equiv \sum_{l=0}^{L-1} \tilde{u}_{n,l} \widetilde{W}_{j-1,\lfloor \frac{n}{2} \rfloor, t-2^{j-1}l \bmod N}, \quad t = 0,\cdots,N-1 \tag{231b}$$

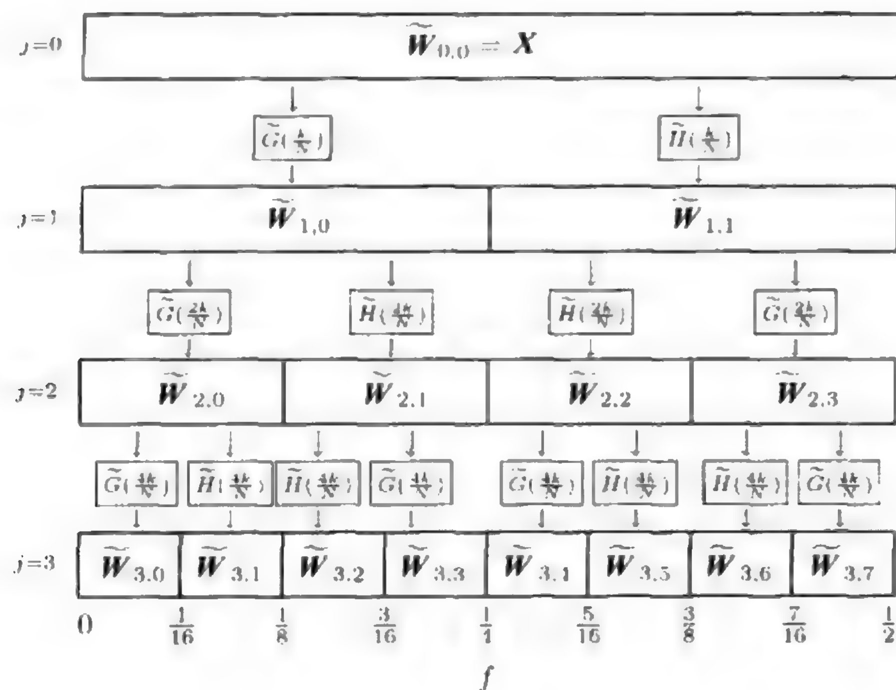


图 232 描述 $\boldsymbol{X}$ 到 $j=1$，2 和 3 层的极大重叠离散小波变换系数 $\widetilde{\boldsymbol{W}}_{j,n}$ 的分析的流程图．共同的，第 j 行的所有系数构成了 $\boldsymbol{X}$ 的 j 层极大重叠离散小波变换．注意在第 j 层内，频率指标 n 从 0 到 2^j-1 范围变化且每个 $\widetilde{\boldsymbol{W}}_{j,n}$ 有长度 N(通过比较，在对离散小波包变换的图 212a 中，每个 $\boldsymbol{W}_{j,n}$ 有长度 $N/2^j$)

(这个陈述产生一个与极大重叠离散小波变换算法十分相似的极大重叠离散小波包变换塔式算法——参见 5.5 节)．我们通过定义第零层系数为时间序列本身来初始化这个方案，即 $W_{0,0,t}\equiv X_t$．对应每个结点(j,n)均有 N 个极大重叠离散小波包变换系数，且将这些系数置成 N 维向量 $\widetilde{\boldsymbol{W}}_{j,n}$(因此 $\widetilde{\boldsymbol{W}}_{0,0}\equiv\boldsymbol{X}$)．如离散小波包变换的情况，我们可以收集到所有的极大重叠离散小波包变换系数直到，譬如，J_0 层且形成一个极大重叠小波包表．图 232 就是 $J_0=3$ 层的这样一个表的例子．

式(231a)和(231b)与对离散小波包变换相对应的公式，即式(215a)和(214a)的比较表明，除了再规范化滤波器的使用，在极大重叠离散小波包变换和离散小波包变换之间最重要的区别是下抽样的消除．因为本质上相同的滤波器被用来产生 $\boldsymbol{W}_{j,n}$ 和 $\widetilde{\boldsymbol{W}}_{j,n}$，故极大重叠离散小波包变换系数与离散小波包变换系数有相同的标称频率区间，即 $\mathcal{I}_{j,n}=\left[\frac{n}{2^{j+1}},\frac{n+1}{2^{j+1}}\right]$．滤波器 $\{\tilde{u}_{j,n,l}\}$ 的传递函数——称为 $\widetilde{U}_{j,n}(\cdot)$——与 $\{u_{j,n,l}\}$ 的传递函数 $U_{j,n}(\cdot)$ 有简单的关系，$\widetilde{U}_{j,n}(f)=U_{j,n}(f)/2^{j/2}$．由于有 $\widetilde{M}_0(f)\equiv M_0(f)/\sqrt{2}=G(f)/\sqrt{2}$ 和 $\widetilde{M}_1(f)\equiv M_1(f)/\sqrt{2}=H(f)/\sqrt{2}$，由式(215b)，我们可以写成

$$\widetilde{U}_{j,n}(f)=\prod_{m=0}^{j-1}\widetilde{M}_{c_{j,n,m}}(2^m f),\tag{232}$$

这里和以前一样，$c_{j,n,m}$ 是二指标值向量 $\boldsymbol{c}_{j,n}$ 的第 m 个元素——这个向量与用在离散小波包变换上的是一样的，且它的结构的例子在图 216 中给出．因为重新尺度

化一个滤波不改变它的相位函数，如果我们在(重新尺度化的)最接近对称或coiflet滤波器基础上构造一个极大重叠离散小波包变换，那么就能够利用在6.5节中发展的提升$|\nu_{j,n}|$来循环平移$\widetilde{W}_{j,n}$的元素，以至于它们——在好的逼近上——与作为X的元素的相同时间有关. 例如，循环平移向量

$$\mathcal{T}^{-|\nu_{j,n}|}\widetilde{W}_{j,n}=[\widetilde{W}_{j,n,|\nu_{j,n}|},\widetilde{W}_{j,n,|\nu_{j,n}|+1},\cdots,\widetilde{W}_{j,n,N-1},\widetilde{W}_{j,n,0},\cdots,\widetilde{W}_{j,n,|\nu_{j,n}|-1}]^{\mathrm{T}}$$

的第t个元素和作为X_t的相同时间有关.

让我们通过证明，以此来总结一下我们的讨论. 任一$j\geqslant1$层的极大重叠离散小波包变换系数都能使用来形成X的能量分解和本质上与基于极大重叠离散小波变换的多分辨分析类似的一个加性分解，这对离散小波包变换来说是正确的(然而，加性分解的每一成分现在都与频率区间$\mathcal{I}_{j,n}$有关而与尺度无关).

• 能量分解

回忆一个时间序列的"能量"是由其元素的平方和给出的，即$\sum_{t=0}^{N-1}X_t^2=\|X\|^2$. 对任一层$j\geqslant1$，我们要求

$$\|X\|^2=\sum_{n=0}^{2^j-1}\|\widetilde{W}_{j,n}\|^2.$$

换句话说，能量保存在j层的极大重叠离散小波包变换里. 为此，令

$$\mathcal{X}_k=\sum_{t=0}^{N-1}X_t\mathrm{e}^{-\mathrm{i}2\pi tk/N},\quad k=0,\cdots,N-1$$

为X的离散傅里叶变换. 因为式(231a)告诉我们，$\widetilde{W}_{j,n}$是用隐含周期化为长度N的滤波器$\{\tilde{u}_{j,n,l}\}$循环卷积X的结果，帕塞瓦尔定理表明

$$\|\widetilde{W}_{j,n}\|^2=\frac{1}{N}\sum_{k=0}^{N-1}\left|\widetilde{U}_{j,n}\left(\frac{k}{N}\right)\right|^2|\mathcal{X}_k|^2$$

(参见式(170c)). 因此，我们有

$$\sum_{n=0}^{2^j-1}\left\|\widetilde{W}_{j,n}\right\|^2=\sum_{n=0}^{2^j-1}\frac{1}{N}\sum_{k=0}^{N-1}\left|\widetilde{U}_{j,n}\left(\frac{k}{N}\right)\right|^2|\mathcal{X}_k|^2=\frac{1}{N}\sum_{k=0}^{N-1}|\mathcal{X}_k|^2\sum_{n=0}^{2^j-1}\left|\widetilde{U}_{j,n}\left(\frac{k}{N}\right)\right|^2.$$

练习[233] 证明对所有的f，有

$$\sum_{n=0}^{2^j-1}|\widetilde{U}_{j,n}(f)|^2=1\text{ 对所有 }f.\qquad(233)$$

◁

因此，从帕塞瓦尔定理的另外的应用中得到结果：

$$\sum_{n=0}^{2^j-1}\|\widetilde{W}_{j,n}\|^2=\frac{1}{N}\sum_{k=0}^{N-1}|\mathcal{X}_k|^2=\sum_{t=0}^{N-1}X_t^2=\|X\|^2.$$

• **加性分解**

在任一层 $j\geqslant 1$ 时，我们要求通过对每个频率带 n 用 $\{\tilde{u}_{j,n,l}\}$ 循环互相关 $\{\widetilde{W}_{j,n,t}\}$ 且求和来重构 $\mathbf{X}$：

$$X_t=\sum_{n=0}^{2^j-1}\sum_{l=0}^{L_j-1}\tilde{u}_{j,n,l}\widetilde{W}_{j,n,t+l\bmod N},\quad t=0,\cdots,N-1. \tag{234}$$

为证明此，我们将证明一个等价的论述是正确的，即将离散傅里叶变换应用到上式，两边形成

$$\mathcal{X}_k=\sum_{n=0}^{2^j-1}\sum_{t=0}^{N-1}\left(\sum_{l=0}^{L_j-1}\tilde{u}_{j,n,l}\widetilde{W}_{j,n,t+l\bmod N}\right)\mathrm{e}^{-\mathrm{i}2\pi tk/N},\quad k=0,\cdots,N-1.$$

现在，互相关的离散傅里叶变换是由周期化为长度为 N 的 $\{\tilde{u}_{j,n,l}\}$ 的离散傅里叶变换的复共轭给出的，即 $\left\{\widetilde{U}^*_{j,n}\left(\frac{k}{N}\right)\right\}$ 用 $\{\widetilde{W}_{j,n,t}\}$ 的离散傅里叶变换相乘，即 $\left\{\widetilde{U}_{j,n}\left(\frac{k}{N}\right)\mathcal{X}_k\right\}$（这可以从式(231a)得出，也就是说 $\{W_{j,n,t}\}$ 是周期化长度为 N 的 $\{\tilde{u}_{j,n,l}\}$ 与 $\{\mathbf{X}_t\}$ 的循环卷积）. 因此，上式等价于

$$\mathcal{X}_k=\sum_{n=0}^{2^j-1}\widetilde{U}^*_{j,n}\left(\frac{k}{N}\right)\widetilde{U}_{j,n}\left(\frac{k}{N}\right)\mathcal{X}_k=\mathcal{X}_k\sum_{n=0}^{2^j-1}\left|\widetilde{U}_{j,n}\left(\frac{k}{N}\right)\right|^2,\quad k=0,\cdots,N-1.$$

然而，式(233)说明右边的和是 1，从而得到了要求的结果.

令 $\widetilde{\mathcal{D}}_{j,n}$ 是包含由循环互相关形成的序列的 N 维向量，并且令 $\widetilde{\mathcal{D}}_{j,n,t}$ 表示它的第 t 个元素. 那么，我们有

$$\widetilde{\mathcal{D}}_{j,n,t}=\sum_{l=0}^{L_j-1}\tilde{u}_{j,n,l}\widetilde{W}_{j,n,t+l\bmod N},\quad \text{因此}\quad X_t=\sum_{n=0}^{2^j-1}\widetilde{\mathcal{D}}_{j,n,t}.$$

我们可以说在时刻 t，值 X_t 是每个频率带 n 里“细节”序列的第 t 项 $\widetilde{\mathcal{D}}_{j,n,t}$ 的和. 因为 $\widetilde{\mathcal{D}}_{j,n}$ 的离散傅里叶变换是由 $\left\{\left|\widetilde{U}_{j,n}\left(\frac{k}{N}\right)\right|^2\mathcal{X}_k\right\}$ 给出，注意这对极大重叠离散小波变换也是正确的，作用在 $\{X_t\}$ 上的复合滤波器的传递函数是由 $|\widetilde{U}_{j,n}(f)|^2$ 给出的，它既是实的又是正的. 因此有零相位. 在时刻 t 的每一频率带里，细节 $\widetilde{\mathcal{D}}_{j,n}$ 与在相同时刻 t 上的时间序列 $\mathbf{X}$ 的特征相一致. 这一有吸引力的特性是未加工的极大重叠离散小波包变换系数 $\widetilde{\mathbf{W}}_{j,n}$ 所不具有的.

6.7 例子：太阳物理数据的极大重叠离散小波包变换

这里我们通过将极大重叠离散小波包变换分析方法应用到太阳物理时间序列上给出了一个例子（这个时间序列在 6.2 节中描述且在图 235 的末端重新绘图）. 为此，我们基于 LA(8)小波滤波器计算 $j=4$ 层的极大重叠离散小波包变换. 考虑到变换是用最好基算法和它的变化选出（参见图 227 和图 228），层的这种选择在时间和频率域上的分辨率之间提供了一个合理的协定：在大约 $\frac{1}{2}$ 周/天以下的频

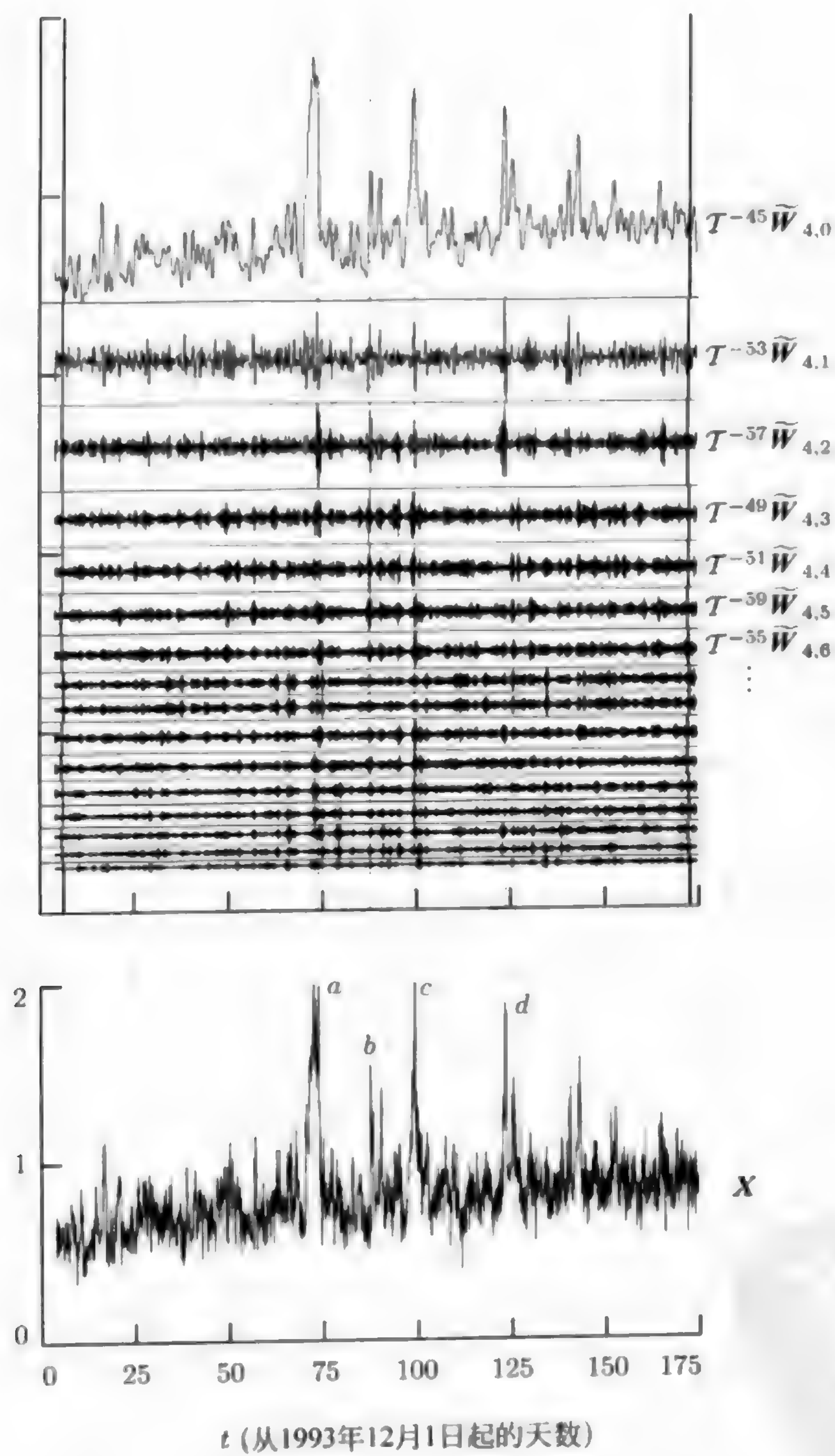


图 235 对太阳物理时间序列的 $j=4$ 层 LA(8) 极大重叠离散小波包变换系数(详情参见正文)

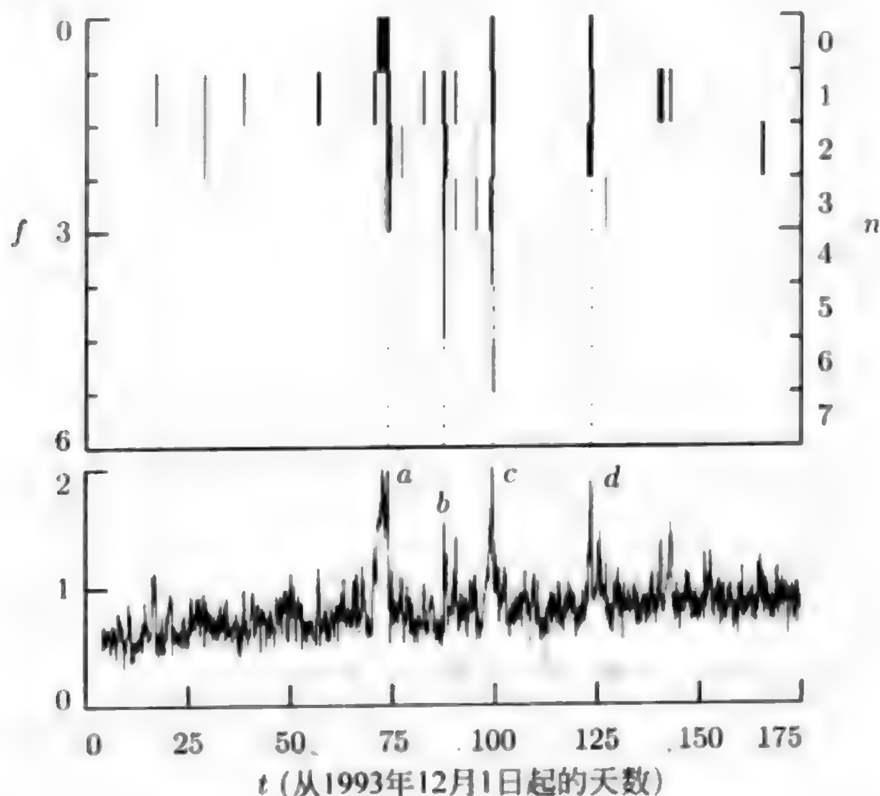


图 236　南半球的太阳磁场量值数据(底图)和相应修改过的时间/频率图(顶图). 这些数据是用毫微特斯拉测量的且由尤里西斯(Ulysses)飞船从 1993 年 338 日(12 月 4 日)世界时间(UT)21 时到 1994 年 144 日(5 月 24 日)世界时间 12 时记录的. 频率 f 是每天的周期. 4 个震动波的结构在底图里标以 a 到 d, 而顶图里的虚线标记了 Balogh 等(1995)所确定的它们的位置

率被著名的太阳旋转结果所支配, 因此这里没有必要为了我们的目的, 很好地解答这个问题, 然而在这个界限以上的频率是由 $j\leqslant 4$ 层的基向量恰当地表示出来的. 一个 $j=4$ 层的变换产生了 16 个向量$\widetilde{\mathbf{W}}_{4,n}$, $n=0$, …, 15. 第 n 个这样的向量对应着频率区间$\left[\frac{n}{2^{j+1}\Delta t}, \frac{n+1}{2^{j+1}\Delta t}\right]$的一个随时间变化的分解. 对 $j=4$ 层的一个极大重叠离散小波包变换, 相关的标准化带宽和时宽分别为$\frac{1}{2^{j+1}}=\frac{1}{32}$和 $\lambda_j=2^j=16$, 其转换成 $1/(2^{j+1}\Delta t)=\frac{3}{4}$周/天和 $\lambda_j\Delta t=\frac{2}{3}$天的实际带宽和时宽. 所有的这些值对分析的目的都是很有意义的.

图 235 的上部表示出了极大重叠离散小波包变换系数向量$\widetilde{\mathbf{W}}_{4,n}$. 在绘出每个向量之前, 我们循环平移它以便它变成近似地由零相位滤波器的输出; 即我们已经绘出了与 X_t 有关的时间的$\mathcal{T}^{-|\nu_{4,n}|}\widetilde{\mathbf{W}}_{4,n}$的第 t 个元素的图, 其中 $\nu_{4,n}$是由 LA(8) 滤波器的相位特性指定的平移(参见式(230)). 对于 $n=0$, …, 6, $|\nu_{4,n}|$ 的实际值在图 235 的右边给出(由于空间限制, $n>7$的值没有全部列出, 因此它们的确定作为习题[6.8]留给读者). 与图 222 一样, 相交于每个$\mathcal{T}^{-|\nu_{4,n}|}\widetilde{\mathbf{W}}_{4,n}$的开头和末端的粗垂线刻画了被循环边界条件影响的极大重叠离散小波包变换系数. 这些

线的设置由一个与极大重叠离散小波变换设置十分相同的论断所确定．这导致了式(198b)，这里，我们可以采用它来告诉我们$\mathcal{T}^{-|\nu_{j,n}|}\widetilde{\boldsymbol{W}}_{j,n}$中边界系数的指标是由

$$t=0,\cdots,L_j-2-|\nu_{j,n}| \text{ 以及 } t=N-|\nu_{j,n}|,\cdots,N-1$$

给出的．虚线标记出了与 Balogh 等(1995)给出的相同的 4 个震荡波的位置．很有意思的是，将这个图里的极大重叠离散小波包变换系数与图 222 里给出的离散小波包变换系数进行比较——由于离散小波包变换里固有的下抽样，所以要看震荡波如何在不同频带里传播是更难的．

图 236 的上部分是一个修改过的时间/频率图，表示出了在$\mathcal{T}^{-|\nu_{4,0}|}\widetilde{\boldsymbol{W}}_{4,0}$中的 100 个最大平方时间前移系数的位置，连同$\mathcal{T}^{-|\nu_{4,1}|}\widetilde{\boldsymbol{W}}_{4,1}$中的 100 个最大的与$\mathcal{T}^{-|\nu_{4,n}|}\widetilde{\boldsymbol{W}}_{4,n}$($n=2$，…，15)中的 100 个最大系数的位置，并把这些作为一组．在这个图中，y 轴左边表示了以物理单位表出的对应每个小波包指标 n(这些在 y 轴右边表示)的频率区间．在$\mathcal{T}^{-|\nu_{4,n}|}\widetilde{\boldsymbol{W}}_{4,n}$($n=2$，…，15)中的 100 个最大平方系数都出现在五个最低的频带内(即标以 $n=2$，…，6 的那些)，因此我们在图 236 中限制频率范围为 0～6 周/天．这个数字很明显地表示出了震荡波特性的时间位置．这里时间平移调整是很重要的．$\mathcal{T}^{-|\nu_{4,n}|}\widetilde{\boldsymbol{W}}_{4,n}$与原时间序列的比较表明在极大重叠离散小波包变换系数中时间局部化的结果出现在与 $\boldsymbol{X}$ 一样的自然时间里且也能容易地穿过频带．当没有时间前移时，结果更难解释(参见练习[6.10])．对于每一个事件，与时间/频率分解的次数一致的选择次数是相关的宽带的(即通过许多频带的延伸)．因为一个严格定义的脉冲有一个宽带的傅里叶分解，故与宽带特性一致的结果位置是很合理的．

在 4 个被标记的事件之中，最有意思的是事件 a．图 236 表示了相对于选择的时间，在宽带特性的发展之前的大约 2.5 天的低频能量(0～0.75 周/天)．一个可能的解释是对于 a，在选择的时间上的事件是一个事件(可能是一个 CIR)，而较早的低频能量来自于另一个事件(可能是 CME)，这是尤里西斯在几乎相同的时间内得出的．这些事件没有必要在太阳系里相同的经度上出现，它们可能是不相似的，因为它们以不同的速度传播给尤里西斯．事件 b 很明显不是多次发生的，大致发生在连接 a 和 c 的太阳圈的半路上．

图 238 的顶部图是标记为 $n=0$，…，15 的频带的细节$\widetilde{\mathcal{D}}_{4,n}$的一个堆积图(包括了 0～12 周/天的整个频率范围)．当细节已被垂直地印刷堆积在一起时，它们都用与时间序列的底部图里用到的相同的相关尺度．震荡波的位置可以通过几个频率带被简单地描绘出．因为这些细节是用零相位滤波器构造的，故这个图帮助确定图 235 和 236 里事件的定时(在这个图和图 235 之间显著的相似性表示 LA(8)滤波器在这里是零相位滤波器的十分好的逼近)．例如，细节序列$\widetilde{\mathcal{D}}_{4,0}$表示了低频能量的存在，是在高频细节$\widetilde{\mathcal{D}}_{4,n}$($n>0$)里的更多特征化的宽带特色之前得到的．(关于这个时间序列的更多细节，参见 Walden and Contreras Cristan，1998a.)

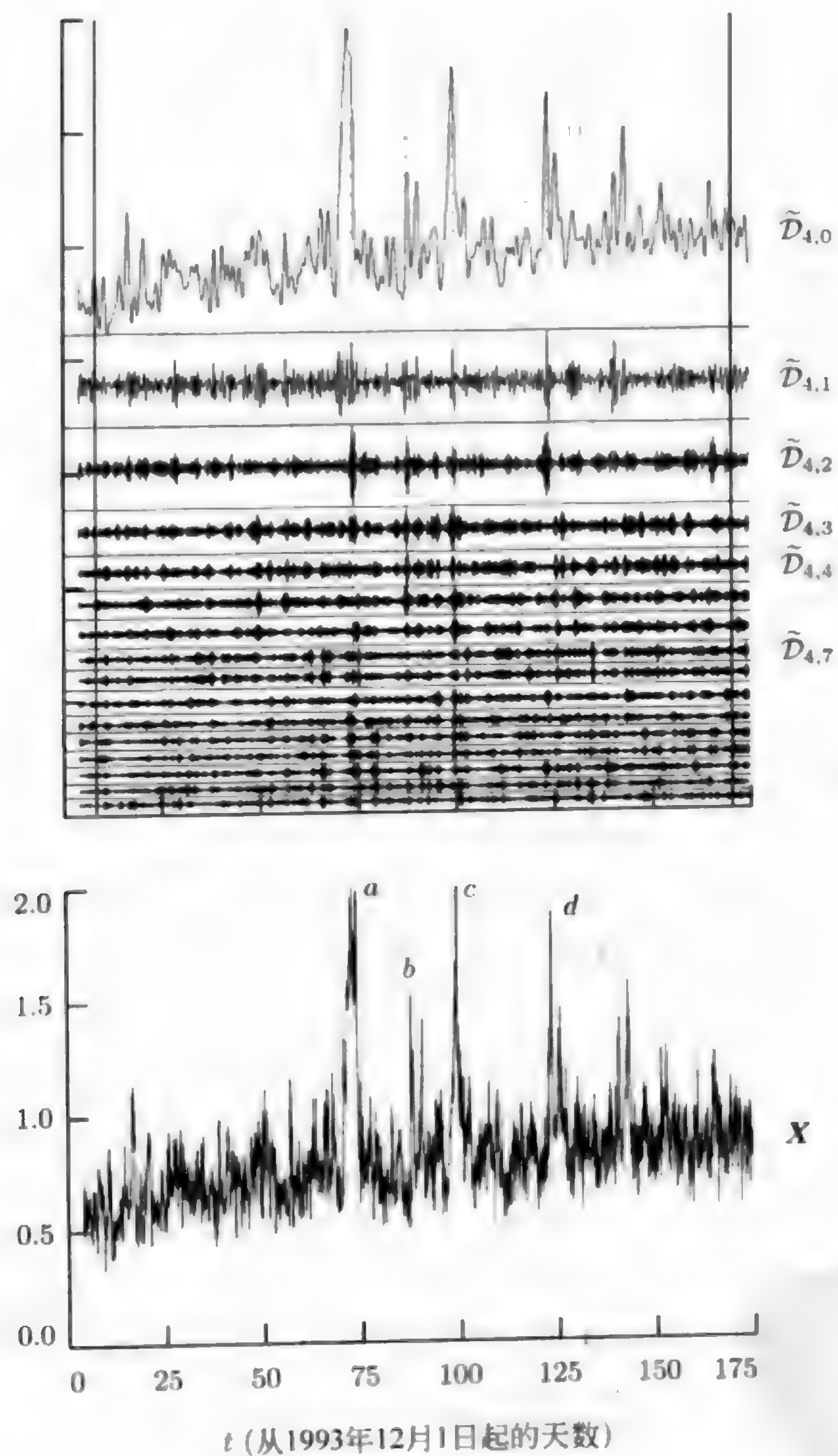


图 238 太阳磁场量值数据(底图)和极大重叠离散小波包变换细节 $\widetilde{D}_{4,n}$，$n=0$，…，15 相应的图(顶图)，第 n 个细节对应了从 $0.75n$ 周/小时到 $0.75(n+1)$ 周/小时的频带

6.8 匹配追踪

匹配追踪背后的原始想法是用比较多的原始函数的一小部分的线性组合来逼近一个函数(Mallat and Zhang，1993；Davis，Mallat and Zhang，1994)．如果我们假定这些原始函数(或波形)的每一个都与一个特定的时间和频率有关，那我们就可以通过利用一个波形和来逼近函数得到它的一个加性的时间/频率分解，而此波形和在时间和频率上的局部化与函数的原始结构相匹配；同样的，我们使用分别与一个特定时间和尺度有关的波形得到一个时间/尺度分解．用于某一目的的从一个大且冗长的集里抽取的波形，称作字典．这里我们在分析离散时间里观察到的时间序列的情况下，考虑这种方法．

和通常一样，设 $\boldsymbol{X}$ 是一个含有时间序列 $\{X_t: t=0, \cdots, N-1\}$ 的元素的 N 维向量．我们希望将 $\boldsymbol{X}$ 扩展到包含一个"字典"$\mathbf{D}$ 里的向量的线性组合：

$$\mathbf{D} \equiv \{\boldsymbol{d}_\gamma : \gamma \in \Gamma\},$$

其中 $\boldsymbol{d}_\gamma = [d_{\gamma,0}, d_{\gamma,1}, \cdots, d_{\gamma,N-1}]^{\mathrm{T}}$ 称为是一个"字典元素"．角标 γ 通常是一个连接 $\boldsymbol{d}_\gamma$ 与时间和频率(或时间和尺度)的参数向量，而 Γ 表示这些参数的所有可能值的集合，且为简单假定是一个有限集(例如，如果 $\boldsymbol{d}_\gamma$ 是离散小波包变换的基向量中的一个，那么 γ 将是一个三维向量 $[j, n, t]^{\mathrm{T}}$，前两个元素将映射到频率，而最后一个元素将映射到时间)．每个字典元素规范化使得

$$\|\boldsymbol{d}_\gamma\|^2 = \sum_{l=0}^{N-1} d_{\gamma,l}^2 = 1.$$

进一步，假设 $\mathbf{D}$ 中元素的某个子集能用来形成 N 维向量空间 $\mathbf{R}^N$ 的一个基，这里假定 $\boldsymbol{X}$ 是空间一个元素．因此，最小的可能字典 $\mathbf{D}$ 是 $\mathbf{R}^N$ 的一个基，但字典是高度冗余的：冗余性允许在构造与 $\boldsymbol{X}$ 的时间/频率或时间/尺度结构十分相匹配的扩展时，有增大的自由度．

匹配追踪是一种算法，用正交映射到 $\mathbf{D}$ 中元素来逼近 $\boldsymbol{X}$．对任一 $\boldsymbol{d}_{\gamma_0} \in \mathbf{D}$，我们将 $\boldsymbol{X}$ 映射到这个向量来形成逼近

$$\boldsymbol{X}^{(1)} \equiv \langle \boldsymbol{X}, \boldsymbol{d}_{\gamma_0} \rangle \boldsymbol{d}_{\gamma_0}$$

(和通常一样，$\langle \cdot, \cdot \rangle$ 表示内积)．然后，我们构造残余向量

$$\boldsymbol{R}^{(1)} \equiv \boldsymbol{X} - \boldsymbol{X}^{(1)} \text{ 所以 } \boldsymbol{X} = \boldsymbol{X}^{(1)} + \boldsymbol{R}^{(1)} = \langle \boldsymbol{X}, \boldsymbol{d}_{\gamma_0} \rangle \boldsymbol{d}_{\gamma_0} + \boldsymbol{R}^{(1)}.$$

下面的练习表明事实上这可以产生一个正交分解．

练习[240] 证明$\langle \boldsymbol{X}^{(1)}, \boldsymbol{R}^{(1)} \rangle = 0$，且因此

$$\|\boldsymbol{X}\|^2 = \|\boldsymbol{X}^{(1)}\|^2 + \|\boldsymbol{R}^{(1)}\|^2 = |\langle \boldsymbol{X}, \boldsymbol{d}_{\gamma_0} \rangle|^2 + \|\boldsymbol{R}^{(1)}\|^2. \quad \triangleleft$$

为了最小化$\|\boldsymbol{R}^{(1)}\|^2$，我们选择$\gamma_0 \in \Gamma$使得

$$|\langle \boldsymbol{X}, \boldsymbol{d}_{\gamma_0} \rangle| = \max_{\gamma \in \Gamma} |\langle \boldsymbol{X}, \boldsymbol{d}_{\gamma} \rangle|.$$

下一步是$\boldsymbol{R}^{(1)}$，将它映射到与$\boldsymbol{R}^{(1)}$匹配得最好的 $\mathbf{D}$ 中的向量：

$$\boldsymbol{R}^{(1)} = \langle \boldsymbol{R}^{(1)}, \boldsymbol{d}_{\gamma_1} \rangle \boldsymbol{d}_{\gamma_1} + \boldsymbol{R}^{(2)},$$

其中$|\langle \boldsymbol{R}^{(1)}, \boldsymbol{d}_{\gamma_1} \rangle| = \max_{\gamma \in \Gamma} |\langle \boldsymbol{R}^{(1)}, \boldsymbol{d}_{\gamma} \rangle|$. 这就导致了分解 $\boldsymbol{X} = \boldsymbol{X}^{(2)} + \boldsymbol{R}^{(2)}$，其中（设$\boldsymbol{R}^{(0)} \equiv \boldsymbol{X}$）

$$\boldsymbol{X}^{(2)} \equiv \boldsymbol{X}^{(1)} + \langle \boldsymbol{R}^{(1)}, \boldsymbol{d}_{\gamma_1} \rangle \boldsymbol{d}_{\gamma_1} = \sum_{k=0}^{1} \langle \boldsymbol{R}^{(k)}, \boldsymbol{d}_{\gamma_k} \rangle \boldsymbol{d}_{\gamma_k}.$$

我们重复这种迭代，经过 m 步后，得到

$$\boldsymbol{X} = \boldsymbol{X}^{(m)} + \boldsymbol{R}^{(m)}, \text{其中 } \boldsymbol{X}^{(m)} \equiv \sum_{n=0}^{m-1} \langle \boldsymbol{R}^{(n)}, \boldsymbol{d}_{\gamma_n} \rangle \boldsymbol{d}_{\gamma_n}. \tag{240}$$

能量保留了，因为

$$\|\boldsymbol{X}\|^2 = \sum_{n=0}^{m-1} \|\langle \boldsymbol{R}^{(n)}, \boldsymbol{d}_{\gamma_n} \rangle \boldsymbol{d}_{\gamma_n}\|^2 + \|\boldsymbol{R}^{(m)}\|^2 = \sum_{n=0}^{m-1} |\langle \boldsymbol{R}^{(n)}, \boldsymbol{d}_{\gamma_n} \rangle|^2 + \|\boldsymbol{R}^{(m)}\|^2.$$

随着 m 的增大，残余能量$\|\boldsymbol{R}^{(m)}\|^2$ 不能增加. 随着 m 的增大，$\|\boldsymbol{R}^{(m)}\|^2$ 减小到 0，这一条件在 Mallat and Zhang(1993)和 Mallat(1998)中讨论过了.

• **离散小波变换和其他正交字典**

作为字典 $\mathbf{D}$ 的第一个例子，我们考虑 J_0 层的部分离散小波变换，为此我们必须假定$N = q2^{J_0}$，其中 J_0 和 q 都是正整数. 和以前一样，令$\mathcal{W}$表示定义这个离散小波变换的 $N \times N$ 阶规范正交矩阵，且设 $\boldsymbol{W} \equiv \mathcal{W}\boldsymbol{X}$ 是离散小波变换系数. 如果我们令$\mathcal{W}_{j\cdot}$是$\mathcal{W}$的第 j 行元素组成的列向量，我们可以写为

$$\boldsymbol{X} = \sum_{j=0}^{N-1} \langle \boldsymbol{X}, \mathcal{W}_{j\cdot} \rangle \mathcal{W}_{j\cdot} = \sum_{j=0}^{N-1} W_j \mathcal{W}_{j\cdot},$$

其中 W_j 是 $\boldsymbol{W}$ 的第 j 个元素. 因为$\mathcal{W}$是一个规范正交变换，我们必定对所有的 j，有$\|\mathcal{W}_{j\cdot}\|^2 = 1$. 离散小波变换字典 $\mathbf{D}^{(\mathrm{dwt})}$ 因此能取为 N 个向量$\mathcal{W}_{j\cdot}$，$j = 0, \cdots, N-1$. $\mathbf{D}^{(\mathrm{dwt})}$ 中的向量形成 $\mathbf{R}^N$ 的一个基，且在 $\mathbf{D}^{(\mathrm{dwt})}$ 中没有冗余. 在第一步，匹配追踪算法只是从 $\mathbf{D}^{(\mathrm{dwt})}$ 挑选出与$|W_j|$对应的最大向量；在 m 步，挑选出了对应第

m 大的向量 $|W_j|$. 这个相同的模式对任何正交变换都保持，因此，对这样的变换，匹配追踪算法只是根据变换系数的大小在向量里增加构成的. 也可以证明，如果我们直到 N 步一直运用匹配追踪算法，就一定有 $\|\boldsymbol{R}^{(N)}\|^2=0$(这是习题[6.11]的部分).

• 小波包表格字典

作为冗余字典的一个例子，我们考虑一个包含了以特定小波滤波器 $\{h_l\}$ 为基础的用来产生直到 J_0 层的小波包表格的所有基向量的字典. 注意基于小波的向量集合与在最好基算法里考虑到的那些是完全一样的. 例如，当 $J_0=2$ 时，这个字典是由包含在 $\mathcal{B}_1^T$, $\mathcal{A}_1^T$, $\mathcal{B}_1^T\mathcal{A}_2^T$, $\mathcal{B}_1^T\mathcal{B}_2^T$, $\mathcal{A}_1^T\mathcal{B}_2^T$ 和 $\mathcal{A}_1^T\mathcal{A}_2^T$ 的 $2N$ 个列向量组成的.

虽然匹配追踪和最好基算法都包括了相同的基向量，但最后的结果是完全不相同的. 再一次用 $J_0=2$ 的例子，最好基算法将根据下列五个规范正交变换中的一个描绘 $\boldsymbol{X}$ 而结束：

$$\mathcal{O}^{(1)}\equiv I_N, \mathcal{O}^{(2)}\equiv\begin{bmatrix}\mathcal{B}_1\\ \mathcal{A}_1\end{bmatrix}, \mathcal{O}^{(3)}\equiv\begin{bmatrix}\mathcal{B}_1\\ \mathcal{B}_2\mathcal{A}_1\\ \mathcal{A}_2\mathcal{A}_1\end{bmatrix}, \mathcal{O}^{(4)}\equiv\begin{bmatrix}\mathcal{A}_2\mathcal{B}_1\\ \mathcal{B}_2\mathcal{B}_1\\ \mathcal{A}_1\end{bmatrix}, \mathcal{O}^{(5)}\equiv\begin{bmatrix}\mathcal{A}_2\mathcal{B}_1\\ \mathcal{B}_2\mathcal{B}_1\\ \mathcal{B}_2\mathcal{A}_1\\ \mathcal{A}_2\mathcal{A}_1\end{bmatrix}.$$

如果被选择的变换是，譬如 $\mathcal{O}^{(5)}$，那么我们就能利用 $\boldsymbol{O}^{(5)}\equiv\mathcal{O}^{(5)}\boldsymbol{X}$ 里的 m 个最大值的系数来逼近 $\boldsymbol{X}$(这与字典里将 $\mathcal{O}^{(5)}$ 运用到匹配追踪算法的 m 步是等价的). 因此，对 $\boldsymbol{X}$ 的逼近将会完全根据 $\mathcal{O}^{(5)}$ 转置里的 N 个向量给出. 相反，如果我们用相同的小波包表实行匹配追踪算法，就将以从 $\mathcal{O}^{(5)}$ 的转置中及从 $\mathcal{O}^{(2)}$ 的转置中选一些向量来结束. 我们现在有两倍多的逼近向量可以选择，这个事实表明对 $\boldsymbol{X}$ 的 m 层的匹配追踪逼近应该比从最好基算法提取的相似逼近有更小的残余(与用 $\|\boldsymbol{R}^{(m)}\|^2$ 衡量的一样). 这点将在 Bruce and Gao(1996a)的 10.4 节中更深入地讨论，且给出了一个匹配追踪确实优于最好基算法的例子.

• 结合的小波包表格和正交离散傅里叶变换字典

这里我们通过将正交离散傅里叶变换(ODFT)里的基向量算到以前的例子里来增大字典，这在 3.4 节中已被讨论. 这些向量的包含开辟了简洁地表示具有宽——或窄——带时间/频率特性的时间序列的可能性. 因为正交离散傅里叶变换是复值的，故如果我们想要逼近 $\boldsymbol{X}^{(m)}$，就需要轻微地修改匹配追踪算法，因此冗余 $\boldsymbol{R}^{(m)}$ 是实值的. 相应地，假设在 $n+1$ 步最大化 $|\langle\boldsymbol{R}^{(n)}, \boldsymbol{d}_\gamma\rangle|$ 的向量 $\boldsymbol{d}_\gamma\in\mathbf{D}$ 是正交离散傅里叶变换向量 $\mathcal{F}_{k\cdot}$，且这个向量是与满足 $0<f_k<1/2$ 的一个频率 $f_k\equiv k/N$ 有关的. 如果我们令 $F_k\equiv\langle\boldsymbol{R}^{(n)}, \mathcal{F}_{k\cdot}\rangle$ 且回忆限制 $F_{N-k}^*=F_k$，我们可看到，如果选择 $\boldsymbol{d}_\gamma=\mathcal{F}_{k\cdot}$ 最大化 $|\langle\boldsymbol{R}^{(n)}, \boldsymbol{d}_\gamma\rangle|$，那么 $\boldsymbol{d}_\gamma=\mathcal{F}_{N-k\cdot}$. 这表明我们应该包含基向量以使式(240)里对逼近

$\mathbf{X}^{(m)}$的第 n 项是由

$$\langle \mathbf{R}^{(n)}, \mathcal{F}_{k\cdot} \rangle \mathcal{F}_{k\cdot} + \langle \mathbf{R}^{(n)}, \mathcal{F}_{N-k\cdot} \rangle \mathcal{F}_{N-k\cdot}$$

给出的，它必然是实值的(参见式(50)和练习[50a])．(注意如果最大化的向量是 $\mathcal{F}_{0\cdot}$ 或 $\mathcal{F}_{\frac{N}{2}\cdot}$，那这个修改就没有必要了，因为这两个都是实值的.)

• **极大重叠离散小波变换字典**

虽然 $\mathbf{X}$ 能根据被小波系数 $\mathbf{W} = \mathcal{W}\,\mathbf{X}$ 加权过的$\mathcal{W}$的(转置)行确切地表示出来，但我们在 5.1 节里已经看到这些系数主要与为时间序列假设的起点有关．而且，我们希望能很容易地处理一个长度不必要是 2 次幂的倍数的序列，因此我们这里考虑用建立在第 5 章里讨论的极大重叠离散小波变换基础上的一个字典 $\mathbf{D}^{(\text{modwt})}$．因此，经过小波滤波器和 J_0 层的选择后，我们定义 $\mathbf{D}^{(\text{modwt})}$ 是$(J_0+1)N$ 个向量的集合，这些向量是由像下边所确定的用一个合适的重新规范化的$\widetilde{\mathcal{W}}_j$，$j=1,\cdots,J_0$ 和$\widetilde{\mathcal{V}}_{J_0}$的行转置给出的．式(171c)表明$\widetilde{\mathcal{W}}_j$ 里的每个行向量，在以长度 N 为周期进行周期化以后是由从极大重叠离散小波变换小波滤波器$\{\tilde{h}_{j,l}\}$里挑选出的元素构成的．现在这个滤波器的宽度由 $L_j\equiv(2^j-1)(L-1)+1$ 给出．如果 $L_j\leqslant N$，那么周期化的滤波器$\{\tilde{h}^{\circ}_{j,l}\}$就与$\{\tilde{h}_{j,l}\}$有相同的非零元素，故我们就可用事实

$$2^j\sum_{j=0}^{L_j-1}\tilde{h}^2_{j,l}=1$$

来再一次规范化$\widetilde{\mathcal{W}}_j$ 的行以得到单位范数——事实上，因为$\tilde{h}_{j,l}\equiv h_{j,l}/2^{j/2}$，这个重新规范化绝对引导我们转向了离散小波变换滤波器．如果我们选择 J_0 满足 $L_{J_0}\leqslant N$，即

$$J_0<\log_2\left(\frac{N}{L-1}+1\right)\text{或者，}\quad \text{等价地，}J_0\leqslant\left\lfloor\log_2\left(\frac{N-1}{L-1}+1\right)\right\rfloor,\qquad(242)$$

那么重新规范化$\widetilde{\mathcal{W}}_j$ 和$\widetilde{\mathcal{V}}_{J_0}$的行就是一个平凡的任务．如果我们选 J_0 满足对某些 $j\leqslant J_0$ 有 $L_j>N$，那么如果 N 是一个顺从 j 层离散小波变换的抽样值，我们仍可用 $2^{j/2}$ 来乘以重新规范化$\widetilde{\mathcal{W}}_j$ 的行(这是因为，经过重新规范化和一个可能的循环平移以后，极大重叠离散小波变换矩阵$\widetilde{\mathcal{W}}_j$ 里的每一行就与离散小波变换矩阵 $\mathcal{W}_j$ 的某行相同，对此，我们有$\mathcal{W}_j\,\mathcal{W}_j^{\mathrm{T}}=I_{N_j}$)．如果 N 不是这样的，那我们仅有的求助就是明确地计算

$$C_j\equiv\sum_{j=0}^{N-1}[\tilde{h}^{\circ}_{j,l}]^2$$

且用$\frac{1}{\sqrt{C_j}}$乘以$\widetilde{\mathcal{W}}_j$ 的行(当 $L_{J_0}>N$ 时，类似的论断对$\widetilde{\mathcal{V}}_{J_0}$仍保持)．

6.9 例子：子潮海平面

这里我们通过用字典 $\mathbf{D}^{(\text{modwt})}$ 来分析 5.8 节里讨论的子潮海平面的涨落来阐述匹配追踪的观点. 这个时间序列(在图 246 的底部被绘出)有 $N=8\ 746$ 个观察值，这些观察值是以 $\Delta t=1/2$ 天间隔开的且覆盖了近 12 个全年. 与第 4 章和第 5 章讨论的一样，离散小波变换和极大重叠离散小波变换假设一个循环时间序列使得，例如，第一个观察值 X_0 被认为是先于最后一个观察值 X_{N-1} 的. 因为子潮海平面涨落有一个强大的年度成分，所以他们被认为——至少在某程度上——作为一个循环序列，如果抽样的数目包含了在这里近似正确的年的一个整数. 抽样尺寸是 2 的乘方而不是 4 的乘方，因此只是层 $J_0=1$ 的部分离散小波变换是可能的离散小波变换的标准的阐述. 极大重叠离散小波变换不要求这样的限制. 另外，因为我们可能又很容易地看到开始和结尾的数据，譬如，是六月而不是一月，故我们的分析对假设的起点相对不灵敏. 给出这样考虑和事实，即子潮海平面涨落是由作用在不同尺度范围上的时间无关现象所影响的，极大重叠离散小波变换是这个序列结构分解的一个合适的架构.

像在 5.8 节中，我们用 LA(8)小波滤波器来进行我们的分析，因此 $L=8$. 假设 J_0 是满足不等式(242)的最大整数，即

$$J_0=\left\lfloor \log_2\left(\frac{N-1}{L-1}+1\right)\right\rfloor=10.$$

这个选择推出 j 层的极大重叠离散小波变换字典向量的非零元素是从离散小波变换滤波器$\{h_{j,l}\}$，$j=1,\cdots,J_0$ 或$\{g_{J_0,l}\}$中挑选出来的(当我们重新规范化极大重叠离散小波变换滤波器以得到单位能量时，这些就是我们所得到的). 此字典因此包括了从$\widetilde{\mathcal{W}}_j$ 和$\widetilde{\mathcal{V}}_{J_0}$的$(J_0+1)N=96\ 206$ 个重新规范化的行. 和通常一样，我们将$\widetilde{\mathcal{W}}_j$ 的行与实际尺度 $\tau_j\Delta t\equiv 2^{j-1}\Delta t$ 上的平均值的变化联系起来且将 $\widetilde{\mathcal{V}}_{J_0}$ 的行与实际尺度 $\lambda_{J_0}\Delta t\equiv 2^{J_0}\Delta t=512$ 天的平均值联系起来.

用匹配追踪算法选择的前 20 个基向量 $\boldsymbol{d}_{\gamma_n}$ 在图 244 和 245 中表示出来了. 在图中向量的高度已被调整使其填满了现存的范围——没有用$\langle \boldsymbol{R}^{(n)}, \boldsymbol{d}_{\gamma_n}\rangle$加权，因为需要在式(240)中形成 $\boldsymbol{X}^{(m)}$；然而，每个向量已通过内积符号被$+1$或-1所乘，以使得很容易地看到在 $\boldsymbol{X}$ 中什么特性能被提取出来. 前 10 个向量都与范围从 $\tau_8\Delta t=64$ 天上的变化到 $\lambda_{10}\Delta t=512$ 天平均数的大尺度有关. 第一个选择是一个大尺度的平均值，它将我们的吸引力转到一个在波动变化范围 1982～1983 上的一个长的向上的增加. 前 20 个选择向量中的 7 个与实际尺度 $\tau_9\Delta t=128$ 天上的变化相联系，这有必要解释一下季节性的变化. 这些中的两个($n=3$ 和 $n=8$)相对于其他的五个被倒置了. 解释 $n=8$ 的倒置是有问题的，因为这个向量明显地利用了循环假设，但$n=3$的倒置与 1981，1983，1985，1987 和 1988 年(n 分别等于 2，1，18，6 和 11)的春天里快速下降相比较，指出了 1984 年春天在涨落中的一个更逐渐下降. 这种异常在图 186 里极大重叠离散小波变换多分辨分析的成分$\widetilde{\mathcal{S}}_7$ 的细致研究中也是很明显的. 与在 Percival and Mofjeld(1997)中讨论的

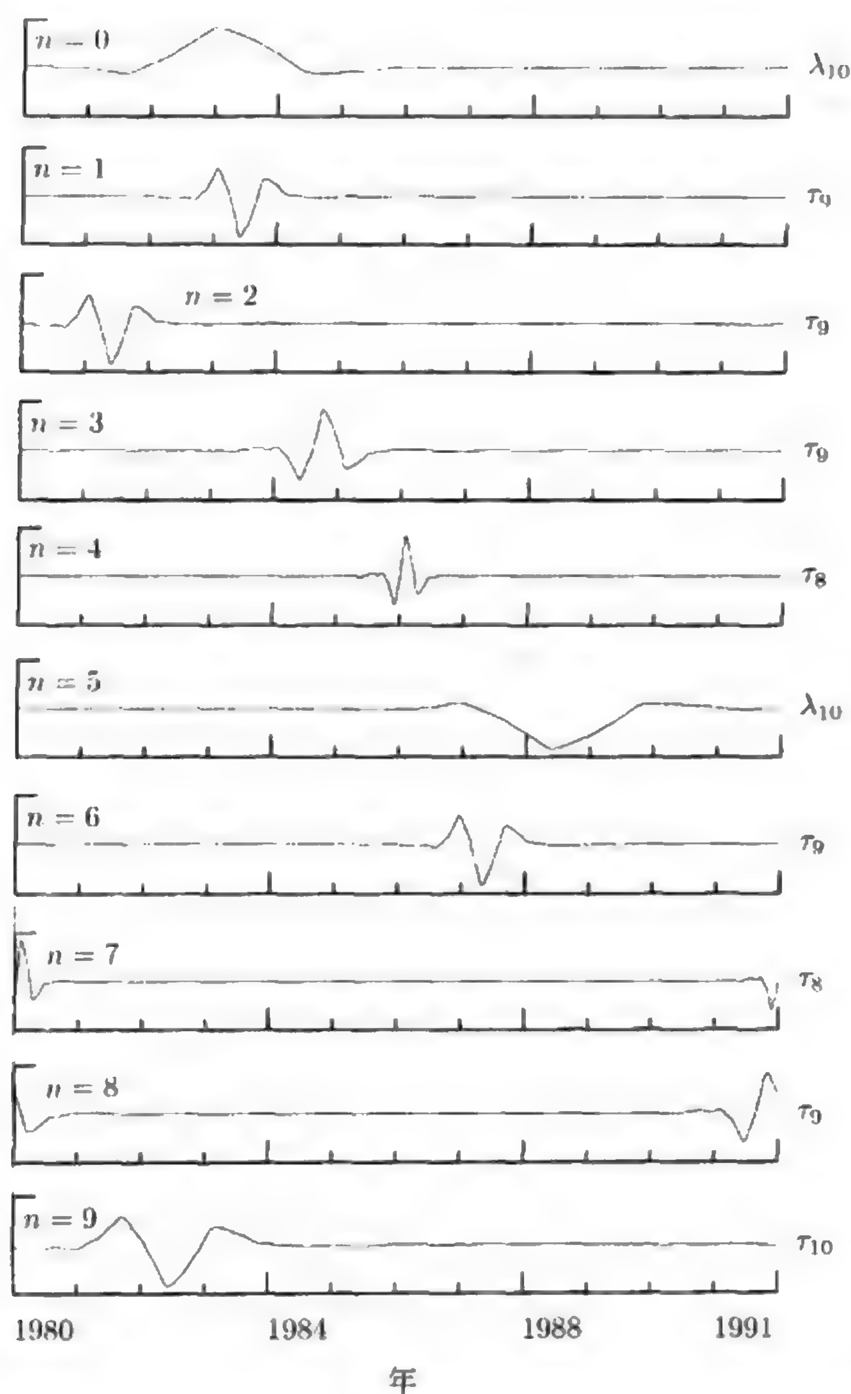


图 244 用匹配追踪算法应用到子潮海平面涨落，选择出的前 10 个极大重叠离散小波变换向量. 字典是用 LA(8)小波滤波器构造的 $\mathbf{D}^{(\text{modwt})}$. 每个向量根据它与 $\boldsymbol{R}^{(n)}$ 的内积符号用 +1 或 −1 来乘，因此能与原序列进行直观相互作用. 每个向量的标准化尺度写在它的图的右边($\widetilde{\mathcal{W}}_j$ 中向量的标准化尺度是 τ_j；而 $\widetilde{\mathcal{V}}_{10}$ 中的是 λ_{10})

一样，3/4 月份海平面上的这些下降是由于北太平洋东边上空气中的春季转变，接着的高压区域海平面是由于低空气压力和向北吹的风. 也注意到，在后 10 个被选择的向量中(图 245)，有 6 个与 $\tau_5\Delta t=8$ 和 $\tau_6\Delta t=16$ 天的尺度上的永久变化有关. 这些都是在冬天或春天里被寻找的且可归因与每年这个时间的暴风雨体系.

图 246 的上部分表示出了前 20 个被选择的向量为基础的对 $\boldsymbol{X}$ 的逼近 $\boldsymbol{X}^{(20)}$

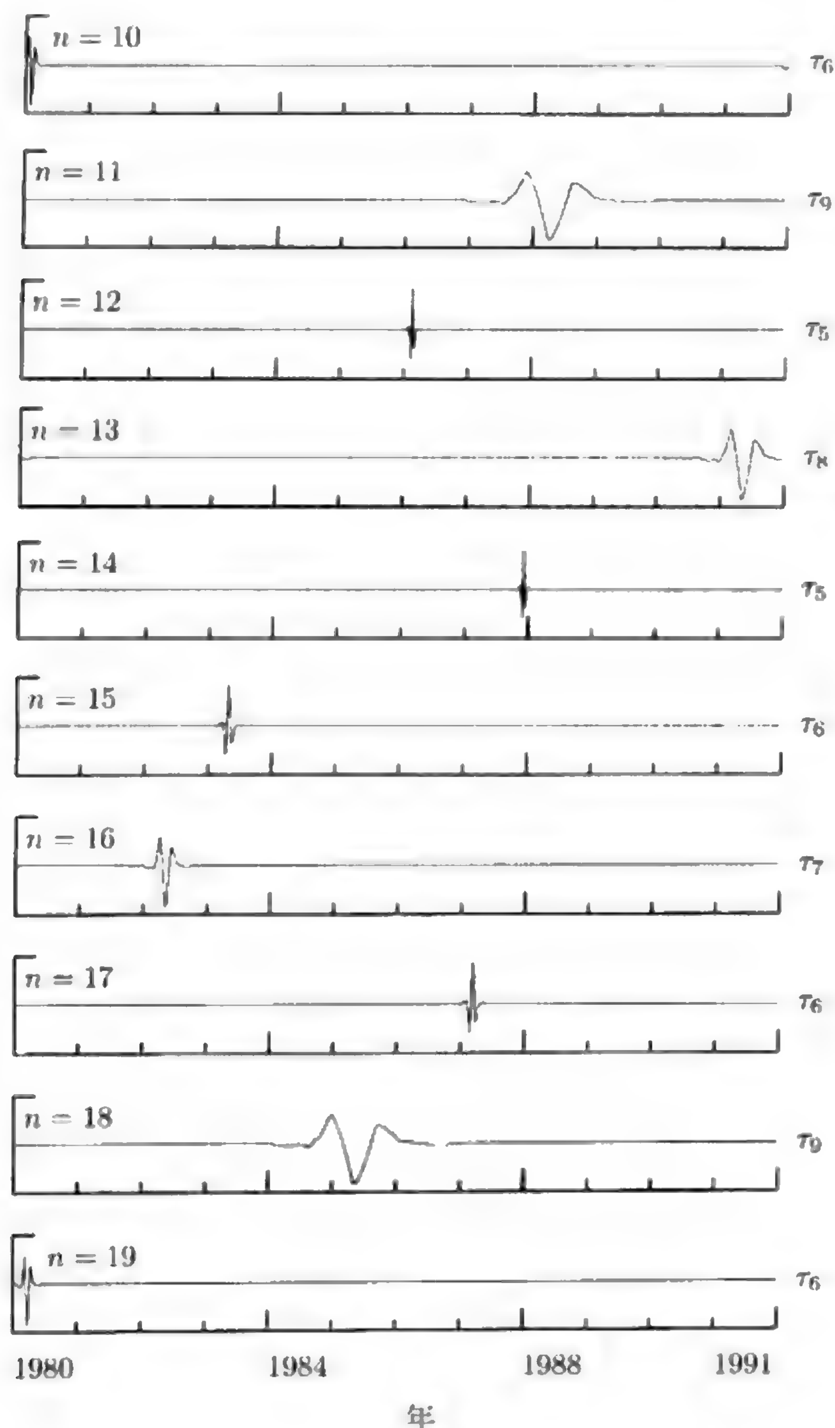


图 245　用匹配追踪算法选出的第 11 到 20 的极大重叠离散小波变换向量(详情参见图 244)

($\boldsymbol{X}$ 本身在图底部表示). 除了特有的 3/4 月份的下降和暴风雨相联系的特性，这个逼近还指出了 1989～1990 的子潮涨落，与这 20 年时间的其他年相比是相对平静的. 如果我们基于 50 个向量升级到一个逼近 $\boldsymbol{X}^{(50)}$，我们就可以看到两年周期里出现的某种结构. 图 246 的最后两个图表示出了 $m=200$ 阶的逼近 $\boldsymbol{X}^{(200)}$ 和相应的剩余 $\boldsymbol{R}^{(200)}$. 注意这些剩余比原序列更均匀(然而，这仍是一个显著的季节相关可变性).

这种逐步的匹配追踪方法清楚地允许我们评论这个序列是怎样构造的，且尤其强调了序列的不同部分构造中的差别. 它传递了与一个使用极大重叠离散小波变换多分辨分析收集到的不同信息(参见 5.8 节).

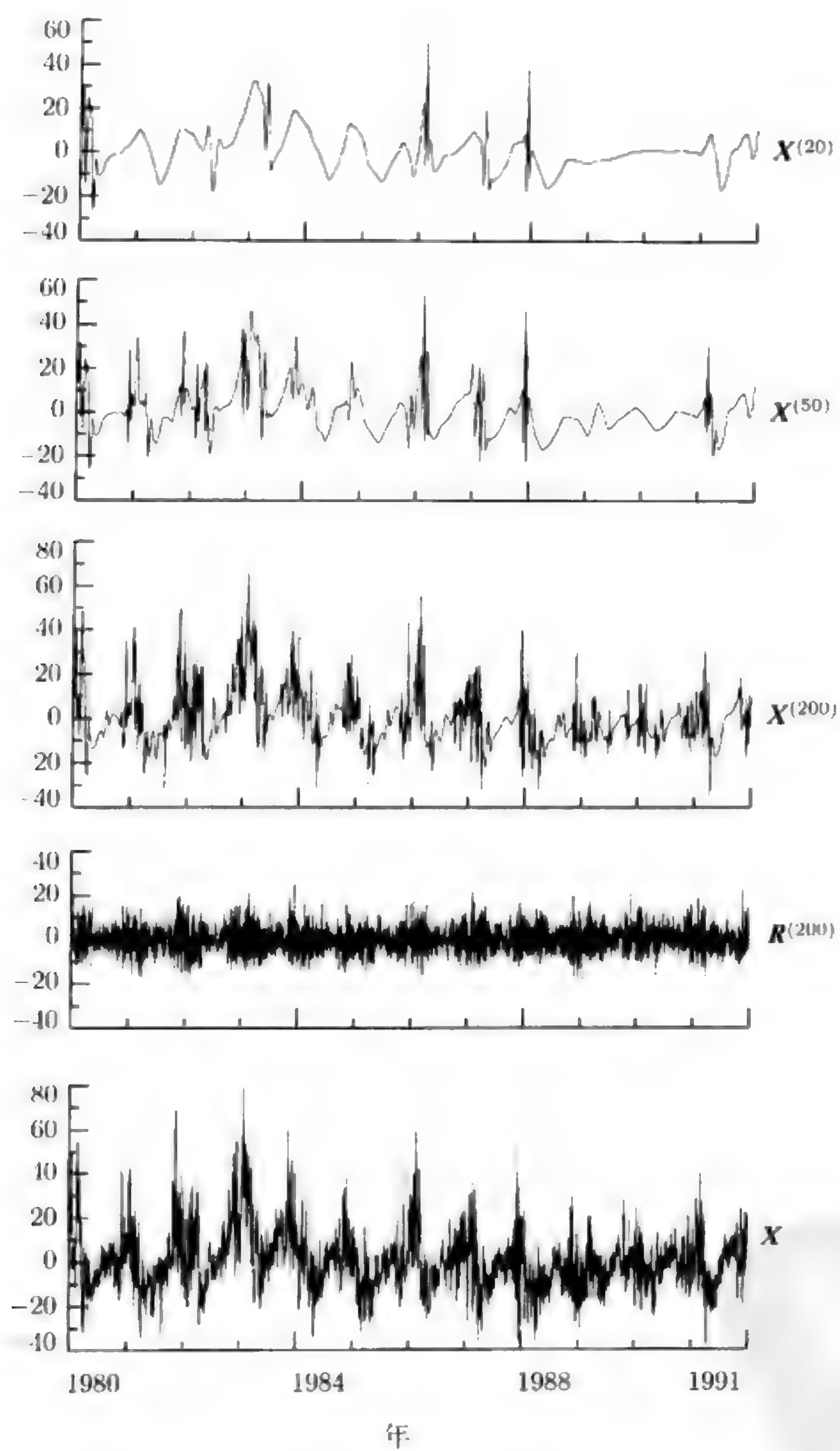


图 246 用极大重叠离散小波变换字典里的 $m=20$，50 和 200 个向量，对子潮海平面涨落 $\boldsymbol{X}$ 的匹配追踪逼近 $\boldsymbol{X}^{(m)}$，相对于 $\boldsymbol{X}^{(200)}$ 的剩余 $\boldsymbol{R}^{(200)}$ 在图 $\boldsymbol{X}$ 的上方表示

6.9 节的评论与扩展

[1]Walden and Contreras Cristan(1998b)为子潮海平面涨落的较短的一段(5年)提供了一个补充的匹配分析. 这个较短的一段的值是关于比我们这里研究的序列的更大参考海平面来测量的. 海平面上的这种变化导致了一个与序列波动相比更大的抽样均值(图 246 中对 $\boldsymbol{X}$ 的抽样均值很小). 作为结果, 他们通过放入一个简单的常数向量来增大字典 $\mathbf{D}^{(\mathrm{modwt})}$, 这个简单常数向量是其中元素均等于 $1/\sqrt{N}$的一个向量. 如果我们假设 $\mathbf{1}$ 是元素均为 1 的一个向量, 那么常数向量就为$\frac{1}{\sqrt{N}}\mathbf{1}$, 所以我们有$\left\langle \boldsymbol{X}, \frac{1}{\sqrt{N}}\mathbf{1}\right\rangle \frac{1}{\sqrt{N}}\mathbf{1}=\overline{X}\mathbf{1}$, 其中 $\overline{X}$ 是 $\boldsymbol{X}$ 的抽样均值. 注意因为这个向量在循环平移下是不变的, 连同 $\mathbf{D}^{(\mathrm{modwt})}$, 使用它产生具有主要特性的一个整体字典, 如果一个向量在这个字典里, 那它所有可能的循环平移也在这个字典里. 常数向量的包含在保持这种主要特性时, 以一种自然灵敏的方式扩大了超定的基. 有了这个增大的字典, Walden and Contreras Cristan(1998b)用匹配追踪算法为它们的比较短的序列求出了第一个向量选择是$\frac{1}{\sqrt{N}}\mathbf{1}$; 即第一个逼近恰好是抽样均值, 且相应的剩余是抽样均值的偏差.

[2]由于在子潮海平面涨落中的突出的每年的变化, 如果考虑一个由 $\mathbf{D}^{(\mathrm{modwt})}$ 结合正交离散傅里叶变换中的向量构成的字典, 我们似乎对这些数据能得到一个更好的匹配追踪逼近. 使用这种增大字典的分析表明, 虽然第一个被选出的向量与图 244 中表示出的向量是一样的, 第二个向量来自正交离散傅里叶变换且有一个于每年一周十分相近的相关周期; 然而, 如果我们直到步 $m=200$ 继续运行匹配追踪算法, 那接下来 198 个向量中只有 3 个是来自正交离散傅里叶变换——剩下所有的都是自 $\mathbf{D}^{(\mathrm{modwt})}$(另 3 个正交离散傅里叶变换向量是在步 $m=65$, 84 和 192 挑选出的且分别与 0 频率, 26.7 周/年和 36.2 周/年相关——这后边的两个可能归因于一到两周的尺度上的天气模型). 对 $m=20$, 50 和 200 的逼近, $\boldsymbol{X}^{(m)}$ 在图 248a 中表示, 且看起来与图 246 中的十分相似——主要差别是年变化在图 248a 中更被完全提高. 另外, 图 248b 表示了利用 $\mathbf{D}^{(\mathrm{modwt})}$(粗曲线)和用正交离散傅里叶变换向量增大的这个字典(细曲线)作为 m 的函数作出的平方的规范化剩余和$\|\boldsymbol{R}^{(m)}\|^2/\|\boldsymbol{X}\|^2$. 虽然当正交离散傅里叶变换向量在第二步中选出时, 我们得到了一个重要的简化, 但逼近 $\mathbf{D}^{(\mathrm{modwt})}$ 接近与步 50 左右的间隔. 因此, 这个序列中特性的大部分比全部频率变动更适合根据局部的时间/尺度变动来描绘.

6.10 小结

令 $\boldsymbol{X}$ 是一个包含实值时间序列的 N 维向量, 此序列的抽样值 N 对某个整数 J_0 来说是 2^{J_0} 的整数倍. 对于 $0\leqslant j\leqslant J_0$, $\boldsymbol{X}$ 的 j 层的离散小波包变换是一个规范

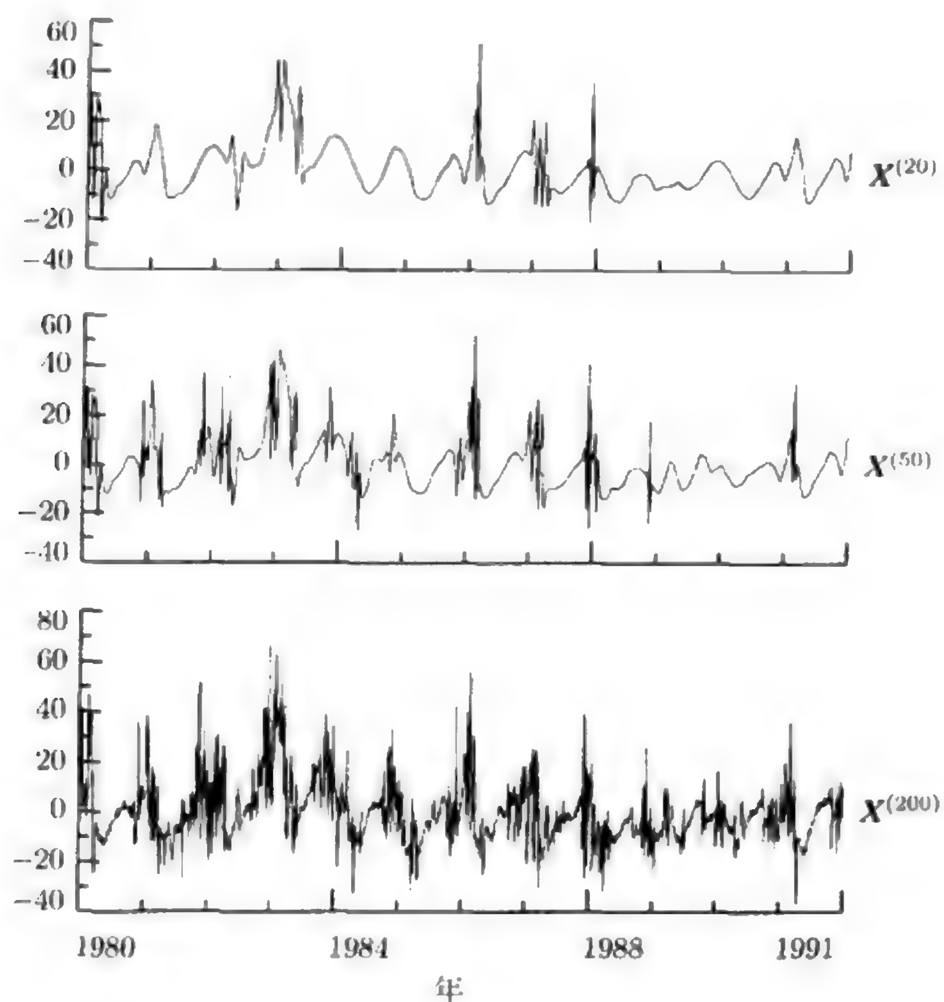


图 248a 使用 $m=20$，50 和 200 个向量，对子潮海平面涨落 $\boldsymbol{X}$ 的匹配追踪逼近 $\boldsymbol{X}^{(m)}$，但现在是从由极大重叠离散小波变换和规范正交离散傅里叶变换向量组成的字典里

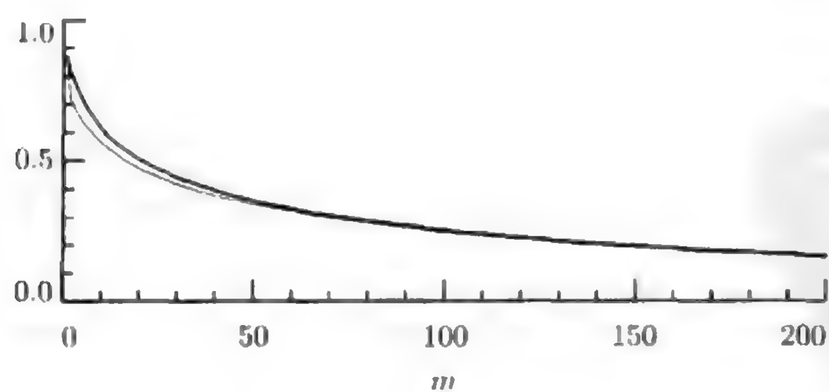


图 248b 在匹配追踪逼近里，使用极大重叠离散小波变换字典 $\mathbf{D}^{(\text{modwt})}$（粗曲线）和这个字典与规范正交离散傅里叶变换的向量结合的字典（细曲线）作为项数 m 的函数的平方的规范化剩余和 $\|\boldsymbol{R}^{(m)}\|^2/\|\boldsymbol{X}\|^2$

正交变换，它能产生一个系数的 N 维向量，能分配为

$$\begin{bmatrix} \mathbf{W}_{j,2^j-1} \\ \mathbf{W}_{j,2^j-2} \\ \vdots \\ \mathbf{W}_{j,1} \\ \mathbf{W}_{j,0} \end{bmatrix},$$

其中每个 $\mathbf{W}_{j,n}$ 均有维数 $N_j \equiv N/2^j$，并且标称上与频率区间 $\mathcal{I}_{j,n} \equiv \left[\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}}\right]$ 有关（这 2^j 个向量一起将标准区间 $\left[0, \frac{1}{2}\right]$ 分划成 2^j 个相等宽度的区间，因此与每个第 j 层离散小波包变换系数相关的带宽是 $1/2^{j+1}$）. 实际上，一个 j 层的离散小波包变换可以用 4.2 和 4.3 节中定义的小波和尺度滤波器 $\{h_l\}$ 和 $\{g_l\}$ 通过滤波一个 $j-1$ 层的离散小波包变换来形成（我们通过定义 $j=0$ 层的离散小波包变换为时间序列本身来初始化这个循环递推方案，即 $\mathbf{W}_{0,0} \equiv \mathbf{X}$）. 如果我们令 $W_{j,n,t}$ 是 $\mathbf{W}_{j,n}$ 的第 t 个元素，那么我们可以写为

$$W_{j,n,t} = \sum_{l=0}^{L-1} u_{n,l} W_{j-1,\lfloor\frac{n}{2}\rfloor, \, 2t+1-l \bmod N_{j-1}}, \quad t = 0, \cdots, N_j - 1.$$

其中

$$u_{n,l} \equiv \begin{cases} g_l & n \bmod 4 = 0 \text{ 或 } 3; \\ h_l & n \bmod 4 = 1 \text{ 或 } 2. \end{cases}$$

并且如果 n 是偶数，那么 $\left\lfloor\frac{n}{2}\right\rfloor$ 就是 $\frac{n}{2}$——否则是 $\frac{n-1}{2}$（上面的公式是式(214a)和(214b)）. 注意从 $j-1$ 层的变换到 j 层的变换时，每个“双亲”向量 $\mathbf{W}_{j-1,n'}$ 被滤波两次（一次用小波滤波器，一次用尺度滤波器），产生两个标记为 $n=2n'$ 和 $n=2n'+1$ 的“孩子”向量 $\mathbf{W}_{j,n}$；然而，小波滤波器当 $n \bmod 4 = 2$ 时产生这些“孩子”向量的第一个（标记为偶数）——如果是 1 时，那就产生第二个“孩子”向量（标记为奇数）. 也注意到，每个孩子都是其双亲长度的一半. 每个孩子中的系数覆盖了原序列 $\mathbf{X}$ 的整个时间范围，但由于小波和尺度滤波器的局部化性质使其在时间上被局部化（这一事实支持了这一论点，即与每个第 j 层离散小波包变换系数相关的时间宽度是 2^j）.

j 层的每个 $\mathbf{W}_{j,n}$ 都是滤波运算独特的有序集的结果，且此运算是应用到层为 $0, 1, \cdots, j-1$ 的系数子集上的. 如果我们用 0 作为尺度滤波器的标号且 1 作为小波滤波器的标号，那我们就能决定一个次序. 按照这个次序，我们需要通过观察长度 j 的二进值向量 $\mathbf{c}_{j,n}$ 的元素应用这些滤波器来构造一个给定的 $\mathbf{W}_{j,n}$，如下构造. 当 $j=1$ 时，我们定义一维向量 $\mathbf{c}_{1,0} \equiv [0]$ 和 $\mathbf{c}_{1,1} \equiv [1]$. 当 $j>1$ 时，我们根据 $j-1$ 层的向量 $\mathbf{c}_{j-1,\lfloor\frac{n}{2}\rfloor}$，通过在 $\mathbf{c}_{j-1,\lfloor\frac{n}{2}\rfloor}$ 末端添加 0（如果 $n \bmod 4 = 0$ 或 3）或者添加 1（如果 $n \bmod 4 = 1$ 或 2）这个简单法则来构造 $\mathbf{c}_{j,n}$. 假定我们已经知道了 $\mathbf{c}_{j-1,\lfloor\frac{n}{2}\rfloor}$——这对于 $j=2$ 是正确的，但对于 $j>2$，我们根据 $j-2$ 层向量再次应用这个简单法则来得到它（如果 $j-2$ 层和更小的是未知的等等）. 作为例子，这个构造步骤产生了 $\mathbf{c}_{4,9} = [1, 1, 0, 1]^T$，它告诉我们先用 $\{h_l\}$ 滤波 $\mathbf{X}$ 以得到 $\mathbf{W}_{1,1}$（经过下抽样）；第二步，用 $\{h_l\}$ 滤波 $\mathbf{W}_{1,1}$ 得到 $\mathbf{W}_{2,2}$；第三步，用 $\{g_l\}$ 滤波 $\mathbf{W}_{2,2}$ 得到 $\mathbf{W}_{3,4}$；最后，用 $\{h_l\}$ 滤波 $\mathbf{W}_{3,4}$ 得到 $\mathbf{W}_{4,9}$.

尽管在实际中，一个 j 层的离散小波包变换通常是根据低级的离散小波包变换循环计算的，但有意思的是 $\mathbf{W}_{j,n}$ 的元素 $W_{j,n,t}$ 能经由式(215a)直接从 $\mathbf{X}$ 得出，即

$$W_{j,n,t} = \sum_{l=0}^{L_j-1} u_{j,n,l} X_{2^j[t+1]-1-l \bmod N}, \quad t = 0,1,\cdots,N_j-1,$$

其中 $\{u_{j,n,l}\}$ 是传递函数由

$$U_{j,n}(f) = \prod_{m=0}^{j-1} M_{c_{j,n,m}}(2^m f)$$

(这就是式(215b))给出的滤波器的脉冲响应序列；这里 $M_0(\cdot) \equiv G(\cdot)$ 和 $M_1(\cdot) \equiv H(\cdot)$ (即 $\{g_l\}$ 和 $\{h_l\}$ 的传递函数，而 $c_{j,n,m}$ 是 $\mathbf{c}_{j,n}$ 的第 m 个元素).

我们也可以根据矩阵运算来阐述离散小波包变换. 因此，我们有

$$\mathbf{W}_{j,n} = \begin{cases} \mathcal{A}_j \mathbf{W}_{j-1,\lfloor \frac{n}{2} \rfloor}, & n \bmod 4 = 0 \text{ 或 } 3; \\ \mathcal{B}_j \mathbf{W}_{j-1,\lfloor \frac{n}{2} \rfloor}, & n \bmod 4 = 1 \text{ 或 } 2. \end{cases}$$

这里 $\mathcal{A}_j$ 和 $\mathcal{B}_j$ 是 $N_j \times N_{j-1}$ 阶矩阵，它们的行包含着分别周期化 $\{g_l\}$ 和 $\{h_l\}$ 为长度 N_{j-1} 所形成的元素(4.12 节概述了怎样构造这些矩阵). 我们可以递推应用上式以得到直接联系 $\mathbf{W}_{j,n}$ 和 $\mathbf{X}$ 的一个矩阵表达式. 为此，我们只需考虑 $\mathbf{c}_{j,n}$ 的元素. 如果 $c_{j,n,m}=0$，那么我们用 $\mathcal{A}_m$ 从 $m-1$ 层到 m 开始；另一方面，如果 $c_{j,n,m}=1$，我们用 $\mathcal{B}_m$. 例如，因为 $\mathbf{c}_{4,9}=[1, 1, 0, 1]^T$，我们可写 $\mathbf{W}_{4,9}=\mathcal{B}_4\mathcal{A}_3\mathcal{B}_2\mathcal{B}_1\mathbf{X}$.

因为离散小波包变换是一个规范正交变换，故我们可以用它由

$$\|\mathbf{X}\|^2 = \sum_{n=0}^{2^j-1} \|\mathbf{W}_{j,n}\|^2$$

分划 $\mathbf{X}$ 的能量，其中 $\|\mathbf{W}_{j,n}\|^2$ 可以被解释为由于带 $\mathcal{I}_{j,n}$ 里的频率引起的对能量的贡献. 类似地，因为 $\mathbf{X}$ 可从离散小波包变换系数综合，故我们可以用 2^j 个向量的加法来表示 $\mathbf{X}$，这其中的每一个向量都是与一个特定频率带 $\mathcal{I}_{j,n}$ 相关的(式(208b)表示出了 $j=2$ 的一个分解).

因为一个 j 层离散小波包变换是通过在 $j-1$ 层离散小波包变换得到每个“母”向量且从中创造出两个“子”向量来形成的，故我们可以将所有 $j=0, 1, \cdots, J_0$ 的离散小波包变换组织成一个小波包表格，对 $J_0=3$ 的一个例子，在图 212a 中表示了出来. 在这个表格里形成离散小波包变换系数向量 $\mathbf{W}_{j,n}$ 的角标数对 (j, n) 可集合在一起形成集 $\mathcal{N} \equiv \{(j, n): j=0, \cdots, J_0, n=0, \cdots, 2^j-1\}$. 每个 $(j, n) \in \mathcal{N}$ 均与一个频率区间，即 $\mathcal{I}_{j,n}$ 有关. 小波包表格可用来形成以非相交二进分解著称的规范正交变换的一个大集合. 通过定义，每个这样的分解都与满足两个性质的属于 $\mathcal{N}$ 的偶子集(譬如 $\mathcal{C}$)有关. 第一，对于每个 $(j, n) \in \mathcal{C}$，所有频率区间 $\mathcal{I}_{j,n}$ 的交集恰恰正是区间 $\left[0, \frac{1}{2}\right]$；第二，如果 (j, n) 和 (j', n') 是 $\mathcal{C}$ 中两个不同的元素，那么 $\mathcal{I}_{j,n}$ 和 $\mathcal{I}_{j',n'}$ 就没有共同的频率. 非相交二进分解包括层 j 的离散小波包变换(对此 $\mathcal{C}=\{(j, 0), (j, 1), \cdots, (j, 2^j-1)\}$)，和 j 层的部分离散小波变换(对此 $\mathcal{C}=\{(j, 0), (j, 1), (j-1, 1), (j-2, 1), \cdots, (1, 1)\}$).

因为如此多的非相交二进分解可以从小波包表格里提取出，故对一给定的时间序列 $\mathbf{X}$，可很有意思地定义一个“最优的”离散小波包表格变换. 最好基算法就是尝试这样做的，它根据形式为

$$M(\mathbf{W}_{j,n}) \equiv \sum_{t=0}^{N_j-1} m(|W_{j,n,t}|)$$

的加性价值泛函分配给表格里每个向量 $\mathbf{W}_{j,n}$ 一个“价值” $M(\mathbf{W}_{j,n})$，这里 $m(\cdot)$ 是定义在 $[0,\infty)$ 上有 $m(0)=0$ 的实值函数. $m(\cdot)$ 的三个例子在 6.3 节里给出. 与在这一节里所讨论的一样，这个算法能选择出非相交二进分解 $\mathcal{C}\in\mathcal{N}$ 使对任何其他非相交二进分解 $\mathcal{C}'\in\mathcal{N}$ 均有

$$\sum_{(j,n)\in\mathcal{C}} M(\mathbf{W}_{j,n}) \leqslant \sum_{(j,n)\in\mathcal{C}'} M(\mathbf{W}_{j,n}).$$

这对离散小波变换是正确的，离散小波包变换对时间序列的假设起点也是灵敏度很高的. 就离散小波变换而言，这种研究导致了极大重叠离散小波变换的定义；对于离散小波包变换，我们导致了类似的“极大重叠”的离散小波包变换（极大重叠离散小波包变换）. 用极大重叠离散小波包变换滤波器

$$\tilde{u}_{n,l} \equiv \frac{u_{n,l}}{\sqrt{2}} \quad \text{和} \quad \tilde{u}_{j,n,l} \equiv \frac{u_{j,n,l}}{2^{j/2}},$$

我们可以定义第 j 层极大重叠离散小波变换系数，其经由

$$\widetilde{W}_{j,n,t} \equiv \sum_{l=0}^{L_j-1} \tilde{u}_{j,n,l} X_{t-l \bmod N}, \quad t=0,1,\cdots,N-1$$

直接根据时间序列 $\mathbf{X}$（这是式(231a)），或者经由

$$\widetilde{W}_{j,n,t} \equiv \sum_{l=0}^{L_j-1} \tilde{u}_{n,l} \widetilde{W}_{j-1,\lfloor\frac{n}{2}\rfloor,t-2^{j-1}l \bmod N}, \quad t=0,\cdots,N-1$$

循环地根据以前确定的 $j-1$ 层系数，其中 $W_{0,0,t}\equiv X_t$（上面的是式(231b). 如果我们将这些系数放在一起成为 N 维向量 $\widetilde{\mathbf{W}}_{j,n}$，那么就能将所有 $j=0,1,\cdots,J_0$ 的极大重叠离散小波包变换组织成一个极大重叠离散小波包变换的小波包表格，对 $J_0=3$ 这样的一个例子，在图 232 中给出.

作为对极大重叠离散小波变换是正确的，极大重叠离散小波包变换对一个时间序列 $\mathbf{X}$ 既导致了能量分解又导致了加性分解. 我们可以经由

$$\|\mathbf{X}\|^2 = \sum_{n=0}^{2^j-1} \|\widetilde{\mathbf{W}}_{j,n}\|^2$$

将 $\mathbf{X}$ 中的能量分划开，这里与离散小波包变换情况相似，$\|\widetilde{\mathbf{W}}_{j,n}\|^2$ 可以被解释为由于带 $\mathcal{I}_{j,n}$ 的频率引起的对能量的贡献. 我们可以用“细节”向量 $\widetilde{\mathcal{D}}_{j,n}$ 形成 $\mathbf{X}$ 的加性分解，它的第 t 个元素是

$$\widetilde{D}_{j,n,t} = \sum_{l=0}^{L_j-1} \tilde{u}_{j,n,l} \widetilde{W}_{j,n,t+l \bmod N}, \text{从这里我们得到 } \mathbf{X} = \sum_{n=0}^{2^j-1} \widetilde{\mathcal{D}}_{j,n}.$$

向量$\widetilde{\mathcal{D}}_{j,n}$与频率区间$\mathcal{I}_{j,n}$有关且根据零相位的滤波器滤波 $\boldsymbol{X}$ 时有一个可选择的阐述——这允许我们简单地将$\widetilde{\mathcal{D}}_{j,n}$里显著的宽带特性与原序列的特性联系起来．极大重叠离散小波包变换系数$\widetilde{\boldsymbol{W}}_{j,n}$本身没有这个性质，但是和 6.5 节里讨论的一样，如果我们用最接近对称或 coiflet 滤波器，然后经由$\mathcal{T}^{-|\nu_{j,n}|}\widetilde{\boldsymbol{W}}_{j,n}$循环提升系数，这里 $\nu_{j,n}$由式(230)给出，我们就能将这些调节到与零相位滤波器的近似输出相适应.

最后，我们得出匹配追踪算法，它被设计用来通过从具有单位范数、大——但有限——的向量集里选出的一小部分向量的线性组合来逼近一个时间序列．这个向量集被称为是一个字典且用 D 表示．尽管在目前章节的上下文里，一个明显的字典将是一个小波包表格(或极大重叠离散小波包变换翻版)，事实上 D 可由种种来源形成，包括——但不被限制——从离散小波变换、极大重叠离散小波变换、离散小波包变换和极大重叠离散小波包变换和规范正交离散傅里叶变换的(可能被重新规范化)向量的各种组合．为了能用 D 中的 m 个向量的线性组合形成对 $\boldsymbol{X}$ 的一种逼近，譬如 $\boldsymbol{X}^{(m)}$，我们随着相关的剩余物 $\boldsymbol{R}^{(n)}$，形成了一系列逼近 $\boldsymbol{X}^{(n)}$，$n=0$，…，$m-1$．逼近 $\boldsymbol{X}^{(n)}$是以 D 中的 n 个向量为基础的(我们定义 $\boldsymbol{X}^{(0)}\equiv\boldsymbol{X}$)，且剩余向量是 $\boldsymbol{X}^{(n)}+\boldsymbol{R}^{(n)}=\boldsymbol{X}$ 这样的．给定第 n 次序的逼近，我们通过计算 $\boldsymbol{R}^{(n)}$ 和每个向量 $\boldsymbol{d}_\gamma\in$ D 之间的内积来得到次序为 $n+1$ 的逼近(角标 γ 用来将一个向量与时间/频率或时间/尺度联系起来——例如，如果 D 包含着与极大重叠离散小波变换相联系的向量，那么 γ 将对 $\boldsymbol{d}_\gamma$ 特殊化尺度和在时间上的位置)．为了形成 $\boldsymbol{X}^{(n+1)}$，我们选择任一向量——称它为 $\boldsymbol{d}_{\gamma_n}$——满足

$$|\langle\boldsymbol{R}^{(n)},\boldsymbol{d}_{\gamma_n}\rangle|\leqslant|\langle\boldsymbol{R}^{(n)},\boldsymbol{d}_\gamma\rangle|,\quad \text{对所有 } \boldsymbol{d}_\gamma\in \mathrm{D}$$

(通常只有一个满足上式的向量，但这里没有只有一个的原因)．新的逼近是这样定义的：

$$\boldsymbol{X}^{(n+1)}\equiv\boldsymbol{X}^{(n)}+\langle\boldsymbol{R}^{(n)},\boldsymbol{d}_{\gamma_n}\rangle\boldsymbol{d}_{\gamma_n},$$

且我们令 $\boldsymbol{R}^{(n+1)}=\boldsymbol{X}-\boldsymbol{X}^{(n+1)}$．经过这样的 m 步后，我们有逼近

$$\boldsymbol{X}^{(m)}=\sum_{n=0}^{m-1}\langle\boldsymbol{R}^{(n)},\boldsymbol{d}_{\gamma_n}\rangle\boldsymbol{d}_{\gamma_n},$$

连同能量分解

$$\|\boldsymbol{X}\|^2=\sum_{n=0}^{m-1}|\langle\boldsymbol{R}^{(n)},\boldsymbol{d}_{\gamma_n}\rangle|^2+\|\boldsymbol{R}^{(m)}\|^2$$

(因为 $|\langle\boldsymbol{R}^{(n)},\boldsymbol{d}_{\gamma_n}\rangle|^2=\|\langle\boldsymbol{R}^{(n)},\boldsymbol{d}_{\gamma_n}\rangle\boldsymbol{d}_{\gamma_n}\|^2$，我们把上式加法的第 n 项看作是与 $\boldsymbol{X}^{(m)}$ 中第 n 个逼近向量有关的能量).

6.11 习题

[6.1] 证明，式(208a)给出的变换是规范正交的.

[6.2] 假设我们在计算直到 J_0 层的离散小波变换时，偶然交换了小波和尺度滤波器$\{h_l\}$和$\{g_l\}$. 借助于 J_0 层的小波包表格，这将会导致什么变换呢？

[6.3] 计算并绘出基于 D(4) 和 LA(16) 滤波器的 $n=0$，…，7 平方增益函数 $|U_{3,n}(\cdot)|^2$ 的图，将这些图与图 217 表示的 LA(8)的情况进行比较.

[6.4] 为了得到在标称频率区间$\left[\frac{19}{128}, \frac{20}{128}\right]$内，表示 $\mathbf{X}$ 里涨落的离散小波包变换系数 $\mathbf{W}_{j,n}$，我们需要对时间序列 $\mathbf{X}$ 实行滤波运算的是什么序列？基于 LA(8)滤波器，计算并绘出相应等价滤波器的脉冲响应序列$\{u_{j,n,l}\}$和平方增益函数$|U_{j,n}(\cdot)|^2$ 的图.

[6.5] 图 216 显示了 $j=1$，2 和 3 层的二进值向量 $\mathbf{c}_{j,n}$的内容. 对于 $j=4$ 和 5 层，确定这些向量.

[6.6] 假设我们计算具有抽样均值 $\bar{X}$ 的一个时间序列 $\mathbf{X}$ 的第 j 层离散小波包变换 $\mathbf{W}_{j,n}$，$n=0$，…，2^j-1. 证明 $\mathbf{W}_{j,0}$的抽样均值等于 $2^{j/2}\bar{X}$. 当 $n>0$ 时，$\mathbf{W}_{j,n}$的抽样均值是 0，这必定是正确的吗？

[6.7] 如图 225 和 226，将最好基算法运用到图 224 的小波包表格中的离散小波包变换系数，但是现在用(a)$p=1$ 的 ℓ_p 信息价值泛函，(b)小波包表格中所有系数绝对值的中间数的阈值集 δ 的阈值价值泛函(包括 $j=0$ 层，即 $\mathbf{W}_{0,0}=\mathbf{X}$；这种设置 δ 的方法在 Bruce and Gao，1996 里用到). 通过这两个价值泛函与用 $-\ell^2\log\ell^2$ 范数所求得的相比较，怎样选择最好基算法？

[6.8] 对于 $j=1$ 到 3，表 230 列出了要求平移滤波器$\{u_{j,n,l}\}$(基于 LA(8)小波滤波器)的前移$|\nu_{j,n}|$. 因此，它有近似零相位. 图 235 用那个图里的标号 $\mathcal{T}^{-|\nu_{4,n}|}\widetilde{\mathbf{W}}_{4,n}$给出了对于 $n=0$，…，6 的$|\nu_{4,n}|$. 对 $j=4$ 层，计算剩下的前移即$|\nu_{j,n}|$，$n=7$，…，15.

[6.9] 类似于图 115 和 124 中对$\{h_{j,l}\}$所做的，绘出滤波器$\{u_{3,n,l}\}$，$n=0$，…，7，的相位函数图；即对这些图中每个基于 Daubechies 或 coiflet 小波滤波器的$\{u_{3,n,l}\}$来说，计算并绘出这个函数被调整前移$|\nu_{3,n}|$以后在标称通带 $\mathcal{I}_{3,n}$ 上确切的相位函数图.

[6.10] 不用前移$|\nu_{4,n}|$，给极大重叠离散小波包变换系数$\widetilde{\mathbf{W}}_{4,n}$绘一个与图 235 类似的图形，怎样将这个图形与图 235 和 238 进行比较？

[6.11] 假设我们有一个长为 $N=2^J$ 的时间序列 $\mathbf{X}$，我们希望用匹配追踪算法来分析. 证明，如果使用任何离散小波变换字典，应用这个算法的 N 步，那么最后一步的剩余 $\mathbf{R}^{(N)}$ 必定满足$\|\mathbf{R}^{(N)}\|^2=0$. 类似地，证明如果用任一基于 J_0 层的极大重叠离散小波变换字典，应用这个算法$(J_0+1)N$ 步，那么最后一步的剩余必定满足$\|\mathbf{R}^{([J_0+1]N)}\|^2=0$.

[6.12] 对 4 个点的时间序列 $\boldsymbol{X}=[2, 4, 8, 10]^{\mathrm{T}}$，基于哈尔小波的 $J_0=2$ 层的极大重叠离散小波变换字典 $\mathbf{D}^{(\mathrm{modwt})}$ 应用匹配追踪算法 4 步. 特别地，对于 $m=1$，2，3 和 4 计算逼近 $\boldsymbol{X}^{(m)}$ 和相对应的剩余 $\boldsymbol{R}^{(m)}$.

[6.13] 使用由向量

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{和} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

组成的字典 $\mathbf{D}$，将匹配追踪算法应用到两个点的时间序列 $\boldsymbol{X}=[5, 3]^{\mathrm{T}}$ 上. 为得到具有零范数的剩余向量 $\boldsymbol{R}^{(m)}$，需要多少步 m?

[6.14] 利用式(224)的 8 个点的时间序列，对于 $j=1$，2 和 3 层的哈尔离散小波包变换，用由所有基向量组成的字典，应用到匹配追踪算法的三步中. 在每一步，此算法能挑选出什么样的向量? 在每一步，$\|\boldsymbol{R}^{(m)}\|^2$ 的值又是多少呢?

第7章

随机变量和随机过程

7.0 引言

在前面3章中，我们导出了离散小波变换和与之相关的变换(包括极大覆盖离散小波变换和离散小波包变换)的基本理论．我们描述了这些理论是怎样应用在时间序列$\{X_t\}$上的，其中到目前为止，我们假设序列中的每个X_t都是实值变量．我们已经证明这些变换产生了十分有用的并且是可加的分解，它可以将序列重新表示成其他一些小的序列，其中每一个都是与有物理意义的独立变量相关的(例如说离散小波变换和极大覆盖离散小波变换情况时的尺度和离散小波包变换情况下的频率区间)．我们还可以使用任何一个变换将时间序列$\{X_t\}$的样本方差分割成与同样的独立变量相关的片断．虽然潜在地十分有用，但是这种可加的分解和对方差的分析事实上是本质的描述，没有考虑到样本方差的重要效果．例如，图192是描述尼罗河年最低水位的序列用极大覆盖离散小波变换得到的那个可加的分解．我们用眼睛可以容易地发现从$\widetilde{\mathcal{D}}_1$和$\widetilde{\mathcal{D}}_2$开始的方差增加．但是，如果我们没有意识到这个序列背后的历史，就会提出-个很好的问题：这种样本方差的增加是不是仅仅是自然的随机的或者是偶然的．为了回答这类问题，我们必须确定时间序列中的随机方差是如何影响小波变换的，并能定量地从变换中计算出来．这就要求我们关于时间序列的观念发生改变：我们现在必须将它看作是一个随机过程$\{X_t\}$的实现(这种观念将在7.4节中定义)．如果我们使用这种过程对时间序列建模，就开始了对基于小波的时间序列分析中样本方差效应的评估任务．

利用统计模型提出时间序列中的问题，这样一种研究方法在上半个世纪中获得了蓬勃的发展(这个主题的一些经典的教科书包括：Anderson，1971；Blackman and Tukey，1958；Bloomfield，1976；Box 等，1994；Brillinger，1981；Fuller，1996；Hannan，1970；Koopmans，1974；Priestley，1981)．在接下来的3章中我们将研究利用小波协助统计模型进行时间序列分析(作为补充，参见 Ogden，1997；Carmona 等，1998；Vidakovic，1999)．为了对这些讨论做

一些准备，这里我们简要地总结概率、统计、随机过程理论中的一些主要结论．我们的目的只是做概要性的叙述，并且只打算限制在接下来的几章中所需要的内容中．对统计和时间序列分析有一定基础的读者可以只是浏览一下本章，主要是熟悉我们书中的符号和定义．希望获得关于统计分析的更多细节的读者可以参阅，例如 Casella and Berger (1990)，Mood 等 (1974)，Papoulis (1991) 或 Priestley(1981)．

7.1 单变量随机变量和概率密度函数

一个实值的随机变量(RV)是一个从随机试验可能结果的样本空间到实轴(或者其子空间)的函数或者映射．令 X 表示一个随机变量，令 x 表示任意可能的输出或者实现，也就是样本空间中的特殊的值．如果 x 可以假设只是一列离散值(例如，整数值)，我们就说 X 是离散随机变量．对这样的随机变量，我们用 $\mathbf{P}[X=x]$ 表示事件 X 取值为 x 的概率．与之相对应的是，如果随机变量取连续的值(例如整个实数轴)，那么这个随机变量就是连续随机变量(在本书中我们主要就是研究这种连续型随机变量)．在这种情况下，谈论事件 $X=x$ 的概率是没有意义的，所以我们的兴趣集中在求事件落在指定区间内的概率，例如 $\mathbf{P}[x<X<x+\Delta]=f_X(x)\Delta$，其中 $f_X(\cdot)$ 是 X 的概率密度函数(PDF)，Δ 是一个无穷小量．概率密度函数有两个基本性质：必定是非负的，并且必定是积分为 1 的．

- *均值和方差*

如果 X 是一个连续的随机变量，它的均值或者期望值是如下定义的常数：

$$E\{X\}=\int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x$$

(这里及在本章任何地方，我们都假设必要量的有限存在性．在这种情况下，我们假定上面的积分存在并且是有限的)．X 的函数 $g(X)$ 的均值是常数

$$E\{g(X)\}=\int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x.$$

X 的方差是由

$$\operatorname{var}\{X\}\equiv E\{(X-E\{X\})^2\}=\int_{-\infty}^{\infty}(x-E\{X\})^2 f_X(x)\,\mathrm{d}x$$

给出的常数．

- *高斯概率密度函数*

最著名的概率密度函数是高斯密度函数(或正态密度函数)，它有如下的形式：

$$f_X(x;\mu,\sigma)=\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)},\quad -\infty<x,\mu<\infty,\sigma>0, \tag{256}$$

其中 μ 和 σ 分别是位置和尺度参数(尽管它们在某些方面有些类似，但是概率密度函数的尺度参数应该与小波滤波中的尺度仔细地区分：前者告诉我们假定的随

机变量的定位参数的值的范围；而后者则告诉我们时间上的有效范围，这些滤波器形成局部加权平均）. 函数 $f_X(\cdot;\mu,\sigma)$ 定义一个对称的钟形曲线（参见图 276 中的粗曲线）. 注意 $f_X(\cdot;\mu,\sigma)$ 常常简写为 $f_X(\cdot)$，也就是说将参数省略掉. 如果 X 是高斯随机变量，我们实际上就有 $E\{X\}=\mu$ 和 $\mathrm{var}\{X\}=\sigma^2$. 可以将“随机变量 X 具有均值为 μ 和方差为 σ^2 的高斯概率密度函数”简记为

$$X \stackrel{\mathrm{d}}{=} \mathcal{N}(\mu,\sigma^2), \tag{257a}$$

这就是说 X 和一个具有参数 μ 和 σ^2 的高斯随机变量具有同样的分布（我们将 $\mathcal{N}(\mu,\sigma^2)$ 解释成这样的随机变量）；上面这个式子读作“X 在分布上等同于 $\mathcal{N}(\mu,\sigma^2)$”. 由 $Z=\frac{X-\mu}{\sigma}$ 定义的随机变量称为标准高斯随机变量或者标准正态随机变量，它具有零均值和单位方差. 要注意的是 $\mathbf{P}[X\leqslant x]=\mathbf{P}[Z\leqslant z]$，其中 $z\equiv\frac{x-\mu}{\sigma}$. 令

$$\Phi(z)\equiv \mathbf{P}[Z\leqslant z]=\int_{-\infty}^{z} f_X(x;0,1)\,\mathrm{d}x=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\,\mathrm{d}x.$$

函数 $\Phi(\cdot)$ 称为高斯累积分布函数，它是 z 的增函数并且对所有有限的 z 满足 $0<\Phi(z)<1$. 这个函数有一个很好定义的反函数 $\Phi^{-1}(\cdot)$，使得对所有的 $0\leqslant p\leqslant 1$ 满足 $\Phi(\Phi^{-1}(p))=p$. 我们有

$$\mathbf{P}[Z\leqslant \Phi^{-1}(p)]=p. \tag{257b}$$

值 $\Phi^{-1}(p)$ 称为标准高斯概率密度函数的 $p\times 100\%$ 百分点，在表 263 中最下面一行中我们选择性地列出一些 p 对应的值. 高斯概率密度函数在下面章节中将会经常使用，尤其是参数 $\mu=0$ 时，此时它是关于原点对称的钟形曲线.

• *一些另外的对称概率密度函数*

拉普拉斯（Laplace）概率密度函数或双边指数概率密度函数关于其局部参数 μ 是对称的，但是是尖状的，并且具有形式

$$f_X(x;\mu,v)=\frac{1}{\sqrt{2/v}}\mathrm{e}^{-(2v)^{1/2}|x-\mu|},\quad -\infty<x,\mu<\infty, v>0, \tag{257c}$$

X 的均值 $E\{X\}=\mu$，方差 $\mathrm{var}\{X\}=1/v$.

高斯概率密度函数和拉普拉斯概率密度函数都属于广义高斯类，它们都有如下形式的概率密度函数：

$$f_X(x;\mu,\alpha,\beta)=\frac{\alpha}{2\beta\Gamma(1/\alpha)}\mathrm{e}^{-(|x-\mu|/\beta)^{\alpha}},\quad -\infty<x,\mu<\infty,\alpha,\beta>0, \tag{257d}$$

其中，$\Gamma(\cdot)$ 是伽马函数，例如参见 Abramowitz and Stegun，1964，第 6 章. 这里 α 是形状参数，并且当 $\alpha=2$ 的时候这个概率密度函数就是高斯概率密度函数，当 $\alpha=1$ 时是拉普拉斯概率密度函数. 广义高斯概率密度函数在均值 μ 处取得峰值，并且随着 α 趋于零而增大.

自由度为 ϑ 的 t 分布的概率密度函数为

$$f_X(x,\vartheta)=\frac{\Gamma([\vartheta+1]/2)}{\Gamma(\vartheta/2)\sqrt{\pi\vartheta}}\left(1+\frac{x^2}{\vartheta}\right)^{-(\vartheta+1)/2},\quad -\infty<x<\infty,\vartheta>0.$$

这个概率密度函数可以进一步推广为包含位置参数μ和尺度参数κ的形式

$$f_X(x,\mu,\kappa,\vartheta)=\frac{\Gamma([\vartheta+1]/2)}{\Gamma(\vartheta/2)\sqrt{\pi\kappa^2\vartheta}}\left(1+\frac{(x-\mu)^2}{\kappa^2\vartheta}\right)^{-(\vartheta+1)/2}, \tag{258a}$$

$-\infty<x$，$\mu<\infty$，$\kappa>0$和$\vartheta>0$. 这个概率密度函数关于μ是对称的并且在μ处平滑地取得峰值点. 如果X是具有这个概率密度函数的随机变量，如果$\vartheta>1$，有$E\{X\}=\mu$；如果$\vartheta>2$，则$\mathrm{var}\{X\}=\kappa^2\vartheta/(\vartheta-2)$. 上述这个分布是所谓的推广了的柯西(Cauchy)分布类的一个特例(例如，参见Johnson等，1994，p. 327). (t和拉普拉斯概率密度函数将在10.4节和10.5节中使用).

• 随机变量的变换

假定$u(\cdot)$是一个可微的单值函数，它的反函数是$v(\cdot)$；即对于所有的x，都有$v(u(x))=x$. 给定随机变量X，假定我们构造$Y=u(X)$，如果X的概率密度函数由$f_X(\cdot)$给出，则Y的概率密度函数为

$$f_Y(y)=|J(y)|\cdot f_X(v(y)), \tag{258b}$$

其中，$J(y)\equiv\frac{\mathrm{d}v(y)}{\mathrm{d}y}$；函数$J(\cdot)$称为(逆)变换的雅可比(Jacobi)行列式. 例如，假定$Y=X^3$，其中X是标准高斯随机变量，则$X\equiv Z$. 这里有$u(x)=x^3$，$v(y)=y^{1/3}$，所以雅可比为$\frac{1}{3}y^{-2/3}$. 我们有

$$f_Y(y)=\frac{1}{3y^{2/3}}f_X(y^{1/3})=\frac{1}{3y^{2/3}\sqrt{2\pi}}\mathrm{e}^{-y^{2/3}/2}.$$

(练习[7.1]给出了应用式(258b)的第二个例子.)

7.2 随机向量和概率密度函数

令X_0和X_1是两个实值的连续随机变量. 令$f_{X_0}(\cdot)$和$f_{X_1}(\cdot)$分别是X_0和X_1的概率密度函数. 记它们在点(x_0, x_1)处的联合概率密度函数为$f_{X_0,X_1}(x_0, x_1)$. 函数$f_{X_0,X_1}(\cdot,\cdot)$在整个实平面(也就是其定义域)上是非负的，并且$f_{X_0,X_1}(\cdot,\cdot)$在这个平面上的二重积分是1.

• 边际概率密度函数和独立性

X_0和X_1的单独的概率密度函数称为是边际密度函数. 每一个边际概率密度函数都可以通过对联合概率密度函数中的另一个变量积分得到：

$$f_{X_0}(x_0)=\int_{-\infty}^{\infty}f_{X_0,X_1}(x_0,x_1)\mathrm{d}x_1;\quad f_{X_1}(x_1)=\int_{-\infty}^{\infty}f_{X_0,X_1}(x_0,x_1)\mathrm{d}x_0.$$

根据定义，如果联合概率函数是边际概率密度函数的乘积，即对任意x_0和x_1均有$f_{X_0,X_1}(x_0, x_1)=f_{X_0}(x_0)f_{X_1}(x_1)$，则随机变量$X_0$和$X_1$是独立的.

• 两个随机变量的函数的期望值

令$g(\cdot,\cdot)$是具有两个自变量的函数，我们用它和X_0及X_1构造新的随机变量，即$g(X_0, X_1)$(一个简单的例子是$g(X_0, X_1)=X_0+X_1$)，则函数$g(X_0, X_1)$的均值或期望值是

$$E\{g(X_0,X_1)\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x_0,x_1)f_{X_0,X_1}(x_0,x_1)\mathrm{d}x_0\mathrm{d}x_1.$$

• 协方差和相关

假设 X_0 和 X_1 的期望值是 $E\{X_0\}=\mu_0$ 和 $E\{X_1\}=\mu_1$. 根据定义，X_0 和 X_1 的协方差是

$$\begin{aligned}\mathrm{cov}\{X_0,X_1\}&\equiv E\{(X_0-\mu_0)(X_1-\mu_1)\}\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_0-\mu_0)(x_1-\mu_1)f_{X_0,X_1}(x_0,x_1)\mathrm{d}x_0\mathrm{d}x_1.\end{aligned}$$

注意 $\mathrm{cov}\{X_0, X_1\}=\mathrm{cov}\{X_1, X_0\}$. 如果我们用这两个随机变量形成一个向量 $\boldsymbol{X}=[X_0, X_1]^{\mathrm{T}}$，$\boldsymbol{X}$ 的期望值和协方差矩阵定义为

$$E\{\boldsymbol{X}\}\equiv\begin{bmatrix}E\{X_0\}\\E\{X_1\}\end{bmatrix},\qquad \Sigma_{\boldsymbol{X}}\equiv\begin{bmatrix}\mathrm{var}\{X_0\} & \mathrm{cov}\{X_0, X_1\}\\ \mathrm{cov}\{X_1, X_0\} & \mathrm{var}\{X_1\}\end{bmatrix},$$

其中，$\mathrm{var}\{X_0\}$是 X_0 的方差：

$$\mathrm{var}\{X_0\}\equiv\mathrm{cov}\{X_0,X_0\}=E\{(X_0-\mu_0)^2\}=\sigma_0^2$$

(类似地，令 $\sigma_1^2\equiv\mathrm{var}\{X_1\}$). 因为 $\mathrm{cov}\{X_0, X_1\}=\mathrm{cov}\{X_1, X_0\}$，所以协方差矩阵是对称的. 随机变量 X_0 和 X_1 的相关度是标准化后的协方差：

$$\rho\equiv\frac{\mathrm{cov}\{X_0,X_1\}}{\sigma_0\sigma_1}.$$

协方差的其他一些性质可以通过练习[7.2]得到.

• 联合高斯随机变量

随机变量 X_0 和 X_1 称为联合高斯(二元高斯)随机变量，如果它们的联合概率密度函数具有如下形式：对 $-\infty<x_0, x_1, \mu_0, \mu_1<\infty$；$\sigma_0, \sigma_1>0$ 和 $-1<\rho<1$，有

$$\begin{aligned}&f_{X_0,X_1}(x_0,x_1;\mu_0,\mu_1,\sigma_0,\sigma_1,\rho)\\&=\frac{1}{2\pi\sigma_0\sigma_1\sqrt{(1-\rho^2)}}\\&\quad\times\exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x_0-\mu_0)^2}{\sigma_0^2}-2\rho\frac{(x_0-\mu_0)(x_1-\mu_1)}{\sigma_0\sigma_1}+\frac{(x_1-\mu_1)^2}{\sigma_1^2}\right]\right).\end{aligned}$$

我们可以有

$$E\{\boldsymbol{X}\}\equiv\begin{bmatrix}\mu_0\\\mu_1\end{bmatrix}$$

和

$$\Sigma_{\boldsymbol{X}}\equiv\begin{bmatrix}\sigma_0^2 & \sigma_0\sigma_1\rho\\ \sigma_0\sigma_1\rho & \sigma_1^2\end{bmatrix}.$$

要注意的是，如果联合高斯随机变量是不相关的，即 $\rho=0$，那么联合密度概率密度函数变成

$$\frac{1}{2\pi\sigma_0\sigma_1}\exp\left(-\frac{(x_0-\mu_0)^2}{2\sigma_0^2}-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}\right)=f_{X_0}(x_0;\mu_0,\sigma_0)f_{X_1}(x_1;\mu_1,\sigma_1),$$

其中，概率密度函数 $f_{X_0}(\cdot;\mu_0,\sigma_0)$和 $f_{X_1}(\cdot;\mu_1,\sigma_1)$由式(256)给出. 因为联合概率密度函数可以写成边际概率密度函数的乘积，所以 X_0 和 X_1 是独立的. (值得注意的是，如果 X_0 和 X_1 不满足联合高斯概率密度，则不相关并不意味着 X_0 和 X_1 是独立的.)

• 条件概率密度函数、均值和中值

另一种定义独立性的方法是借助于条件概率密度函数. 给定 $X_1=x_1$ 后随机变量 X_0 的条件概率密度函数定义为

$$f_{X_0|X_1=x_1}(x_0)=\frac{f_{X_0,X_1}(x_0,x_1)}{f_{X_1}(x_1)}. \tag{260}$$

符号"|"应该读作"给定"(或"在……的条件下"). 所以 $f_{X_0|X_1=x_1}(\cdot)$定义了给定 $X_1=x_1$ 的条件下 X_0 的概率密度函数. 利用式(260)有: X_0 和 X_1 的独立性意味着 $f_{X_0|X_1=x_1}(x_0)\equiv f_{X_0}(x_0)$和 $f_{X_1|X_0=x_0}(x_1)\equiv f_{X_1}(x_1)$，也就是说一个随机变量的取值不影响另一个随机变量的分布. 要注意的是，二元概率密度函数可以通过边际概率密度和条件概率密度函数获得：

$$f_{X_0,X_1}(x_0,x_1)=f_{X_0|X_1=x_1}(x_0)f_{X_1}(x_1),$$

或者根据对称性，有

$$f_{X_0,X_1}(x_0,x_1)=f_{X_1|X_0=x_0}(x_1)f_{X_0}(x_0).$$

与条件分布的概念相联系的是条件均值或条件期望(以及条件方差). 给定 $X_1=x_1$，X_0 的条件均值是

$$E_{X_0|X_1}\{X_0\mid X_1=x_1\}=\int_{-\infty}^{\infty}x_0f_{X_0|X_1=x_1}(x_0)\mathrm{d}\,x_0.$$

同样可以得到给定 $X_0=x_0$ 的 X_1 的条件期望. 可以证明

$$E_{X_1}\{E_{X_0|X_1}\{X_0\mid X_1\}\}=E\{X_0\}$$

和

$$E_{X_0}\{E_{X_1|X_0}\{X_1\mid X_0\}\}=E\{X_1\},$$

即，如果我们求条件随机变量值的条件均值的平均值，就可以得到(无条件)均值. 当条件均值很明显的时候，我们常常将 $E_{X_0|X_1}\{X_0\mid X_1=x_1\}$写作 $E\{X_0\mid X_1=x_1\}$.

下面是条件均值的一个重要的应用. 令 X_0 和 X_1 是两个相联系的随机变量，并且只能观测一个，如 X_0. 假设我们希望通过 X_1 对 X_0 的依赖关系来逼近不可观测的随机变量 X_1. 令 $U_2(X_0)$表示一个新的随机变量，其中 $U_2(\cdot)$是某个非线性函数. 假设我们用均方差 $E\{(X_1-U_2(X_0))^2\}$来度量不可观测的 X_1 和 $U_2(X_0)$的逼近程度. 我们可以断言：当 $U_2(X_0)=E\{X_1\mid X_0\}$(也就是给定 X_0 后 X_1 的条件期望)的时候，$E\{(X_1-U_2(X_0))^2\}$取最小值. 为了看清楚这一点，注意

$$E\{(X_1-U_2(X_0))^2\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-U_2(x_0))^2f_{X_0,X_1}(x_0,x_1)\mathrm{d}x_0\mathrm{d}x_1;$$

然而，$f_{X_0,X_1}(x_0,\ x_1)=f_{X_1|X_0=x_0}(x_1)f_{X_0}(x_0)$，所以有

$$E\{(X_1-U_2(X_0))^2\}=\int_{-\infty}^{\infty}f_{X_0}(x_0)\left[\int_{-\infty}^{\infty}(x_1-U_2(x_0))^2f_{X_1|X_0=x_0}(x_1)\mathrm{d}x_1\right]\mathrm{d}x_0.$$

和 $f_{X_0}(\cdot)$一样，方括号中的积分是非负的，所以上述期望达到最小值当且仅当对每一个可能的结果 x_0，积分 $\int(x_1-U_2(x_0))^2f_{X_1|X_0=x_0}(x_1)\mathrm{d}x_1$ 都尽可能地小. 对给定的 x_0，这个积分是条件概率密度函数 $f_{X_1|X_0=x_0}(\cdot)$关于常数 $U_2(x_0)$的二阶矩.

练习[261] 假设 X 是随机变量，它的概率密度函数是 $f_X(\cdot)$. 如果 a 是常数，证明

$$E\{(X-a)^2\} = \int_{-\infty}^{\infty}(x-a)^2 f_X(x)\mathrm{d}x$$

当 $a=E\{X\}$的时候取得最小值. ◁

所以，用常数逼近一个随机变量，当这个常数等于随机变量的期望的时候，均方误差取得最小值(物理解释是，重心的惯量矩要比其他点的惯量矩都要小). 练习[261]的结果意味着

$$\int(x_1 - E\{X_1 \mid X_0 = x_0\})^2 f_{X_1 \mid X_0 = x_0}(x_1)\mathrm{d}x_1 \leqslant \int(x_1 - U_2(x_0))^2 f_{X_1 \mid X_0 = x_0}(x_1)\mathrm{d}x_1,$$

所以，正如我们所断言的那样，选择 $U_2(x_0)=E\{X_1 \mid X_0 = x_0\}$，即 $U_2(X_0)=E\{X_1 \mid X_0\}$，$E\{(X_1-U_2(X_0))^2\}$ 取得最小值. 为了进一步使用，我们将 $E\{X_1 \mid X_0=x_0\}$是关于 $U_2(x_0)$的如上所述的积分取得极小值写成如下的形式：

$$E\{X_1 \mid X_0 = x_0\} = \arg\min_{U_2(x_0)}\int(x_1 - U_2(x_0))^2 f_{X_1 \mid X_0 = x_0}(x_1)\mathrm{d}x_1. \quad (261a)$$

条件均值估计量就是在最小化均方误差的时候出现的. 我们还可以选择最小化绝对平均误差 $E\{\mid X_1-U_1(X_0)\mid\}$. 在这种情况下，估计量 $U_1(X_0)$是条件中值而不是条件均值(参见，例如 Van Trees，1968，p. 56). 当 $X_0=x_0$ 的时候，我们需要求解

$$\int_{-\infty}^{U_1(x_0)} f_{X_1 \mid X_0 = x_0}(x_1)\mathrm{d}x_1 = 0.5 \quad (261b)$$

来确定 $U_1(x_0)$.

• **随机变量的变换**

以处理依赖于一个以上的随机变量的变换，这里我们推广式(258b). 假设 $u_0(\cdot,\cdot)$和 $u_1(\cdot,\cdot)$是两个自变量的可微单值函数. 这就推出，如果 $y_0=u_0(x_0, x_1)$和 $y_1=u_1(x_0, x_1)$，则存在反函数 $v_0(\cdot,\cdot)$和 $v_1(\cdot,\cdot)$，使得$x_0=v_0(y_0, y_1)$和 $x_1=v_1(y_0, y_1)$. 假设我们使用随机变量 X_0 和 X_1 来创建随机变量的变换

$$Y_0 = u_0(X_0,X_1), \quad Y_1 = u_1(X_0,X_1).$$

如果(逆)变换的雅可比 $J(\cdot,\cdot)$是非零的，则我们可以容易地根据 X_0 和 X_1 的联合概率密度函数 $f_{X_0,X_1}(\cdot,\cdot)$得出 Y_0 和 Y_1 的联合概率密度函数 $f_{Y_0,Y_1}(\cdot,\cdot)$. 因此，如果

$$J(y_0,y_1) \equiv \begin{vmatrix} \dfrac{\partial v_0(y_0,y_1)}{\partial y_0} & \dfrac{\partial v_0(y_0,y_1)}{\partial y_1} \\ \dfrac{\partial v_1(y_0,y_1)}{\partial y_0} & \dfrac{\partial v_1(y_0,y_1)}{\partial y_1} \end{vmatrix}$$

$$= \frac{\partial v_0(y_0,y_1)}{\partial y_0}\cdot\frac{\partial v_1(y_0,y_1)}{\partial y_1} - \frac{\partial v_0(y_0,y_1)}{\partial y_1}\cdot\frac{\partial v_1(y_0,y_1)}{\partial y_0}$$

$$\neq 0,$$

则

$$f_{Y_0,Y_1}(y_0,y_1)=|J(y_0,y_1)|\cdot f_{X_0,X_1}(v_0(y_0,y_1),v_1(y_0,y_1)).$$

• *和的概率密度函数*

现在让我们利用上面的结果来求两个随机变量的和的概率密度函数的表示式.

练习[262a] 证明，如果 $Y_0=X_0$ 和 $Y_1=X_0+X_1$，则

$$f_{Y_0,Y_1}(y_0,y_1)=f_{X_0,X_1}(y_0,y_1-y_0).$$ ◁

我们要求的概率密度函数 $f_{Y_1}(\cdot)$恰恰是上述联合概率密度函数关于和 Y_1 的边际概率密度函数. 因此

$$f_{Y_1}(y_1)=\int_{-\infty}^{\infty}f_{Y_0,Y_1}(y_0,y_1)\mathrm{d}y_0=\int_{-\infty}^{\infty}f_{X_0,X_1}(y_0,y_1-y_0)\mathrm{d}y_0.$$

另外，如果 X_0 和 X_1 是独立的，则 $f_{X_0,X_1}(y_0,\ y_1-y_0)=f_{X_0}(y_0)f_{X_1}(y_1-y_0)$，所以，我们有

$$f_{Y_1}(y_1)=\int_{-\infty}^{\infty}f_{X_0}(y_0)f_{X_1}(y_1-y_0)\mathrm{d}y_0. \tag{262a}$$

即，如果两个随机变量是独立的，则它们和的概率密度函数是它们概率密度函数的卷积.

• *多元随机向量*

二元随机向量的结果可以自然地推广到高维随机向量 $\boldsymbol{X}=[X_0,\ X_1,\ X_2,\ \cdots,\ X_{N-1}]^T$，我们介绍几个我们感兴趣的结果.

令$\mathcal{M}$是一个 $M\times N$ 阶矩阵. 则 $\boldsymbol{Y}=\mathcal{M}\boldsymbol{X}$ 的均值是

$$\mu_{\boldsymbol{Y}}=E\{\boldsymbol{Y}\}=\mathcal{M}E\{\boldsymbol{X}\}=\mathcal{M}\mu_{\boldsymbol{X}}, \tag{262b}$$

它的协方差矩阵是

$$\Sigma_{\boldsymbol{Y}}=\mathcal{M}\Sigma_{\boldsymbol{X}}\mathcal{M}^T. \tag{262c}$$

练习[262b] 假设$\mathcal{M}=\mathcal{O}$，其中$\mathcal{O}$是规范正交变换(因此是 $N\times N$ 阶的). 证明在下述意义下全方差是不变的：

$$\sum_{t=0}^{N-1}\mathrm{var}\{Y_t\}=\sum_{t=0}^{N-1}\mathrm{var}\{X_t\}.$$ ◁

另外，如果我们假设 $\boldsymbol{X}$ 中的 N 维随机变量是多维高斯型随机变量，则 M 维随机变量$\boldsymbol{Y}=\mathcal{M}\boldsymbol{X}$ 也是多维高斯型随机变量. 因为对于高斯型随机变量不相关意味着独立(反之也成立)，则下面的练习表明一列独立同分布(IID)的随机变量有零均值的高斯型随机变量的规范正交变换产生了一列新的独立同分布的随机变量，并且具有同样的联合分布.

表 263

η	p					
	0.005	0.025	0.05	0.95	0.975	0.995
1	0.000 04	0.001 0	0.003 9	3.841 5	5.023 9	7.879 4
1.5	0.001 5	0.013 1	0.033 2	4.980 2	6.275 8	9.331 0
2	0.010 0	0.050 6	0.102 6	5.991 5	7.377 8	10.596 6
2.5	0.032 1	0.118 6	0.210 8	6.928 1	8.392 3	11.753 8
3	0.071 7	0.215 8	0.351 8	7.814 7	9.348 4	12.838 2
3.5	0.130 1	0.338 9	0.520 1	8.665 1	10.262 1	13.869 6
4	0.207 0	0.484 4	0.710 7	9.487 7	11.143 3	14.860 3
4.5	0.301 3	0.649 4	0.920 1	10.288 2	11.998 5	15.818 3
5	0.411 7	0.831 2	1.145 5	11.070 5	12.832 5	16.749 6
5.5	0.537 0	1.027 8	1.384 5	11.837 6	13.648 6	17.658 3
6	0.675 7	1.237 3	1.635 4	12.591 6	14.449 4	18.547 6
6.5	0.826 8	1.458 4	1.896 7	13.334 3	15.236 9	19.420 1
7	0.989 3	1.689 9	2.167 3	14.067 1	16.012 8	20.277 7
7.5	1.162 1	1.930 6	2.446 3	14.791 2	16.778 3	21.122 2
$\Phi^{-1}(p)$	−2.575 8	−1.960 0	−1.644 9	1.644 9	1.960 0	2.575 8

注：χ^2_η 分布的百分点 $Q_\eta(p)$，以步长为 0.5 将 η 从 1 取到 7.5．最下面一行给出了标准高斯分布的百分点 $\Phi^{-1}(p)$．

练习[263] 假设$\mathcal{O}$是规范正交变换．证明：如果 $\boldsymbol{X}$ 是独立同分布有零均值(也就是 $\mu_X=\mathbf{0}$，$\mathbf{0}$ 是零向量)的高斯型随机向量并且协方差是 $\sigma^2 I_N$，则 $\boldsymbol{Y}\equiv\mathcal{O}\,\boldsymbol{X}$ 同样是独立同分布零均值的高斯型随机向量，协方差也是 $\sigma^2 I_N$．也就是说 $\boldsymbol{Y}$ 和 $\boldsymbol{X}$ 的统计特性是一样的．这个结论对 $\mu_X\neq\mathbf{0}$ 成立吗？(你可以假设上面的结论是成立的，也就是说 $\boldsymbol{X}$ 中的随机变量是多元高斯型随机变量，然后证明 $\boldsymbol{Y}$ 中的随机变量也是如此.) ◁

• χ^2 随机变量

下面我们总结一下 χ^2 随机变量的性质，它可以通过独立同分布标准高斯分布的随机变量 Z_1，Z_2，…，Z_η 的平方和来引入，即

$$\sum_{k=1}^{\eta} Z_k^2 \equiv \chi^2_\eta. \tag{263a}$$

随机变量 χ^2_η 称为是服从具有 η 个自由度的 χ^2 分布．χ^2_η 的概率密度函数具有如下形式：

$$f_{\chi^2_\eta}(x;\eta)=\frac{1}{2^{\eta/2}\Gamma(\eta/2)}x^{(\eta/2)-1}\mathrm{e}^{-x/2},\quad x\geqslant 0,\eta>0 \tag{263b}$$

(证明参见习题[7.3])．对所有的 $\eta>0$，上面的函数都是非负的并且积分为 1，所以我们可以通过它来扩展 χ^2 随机变量的定义导出具有不同于非负整数的自由度的 χ^2 随机变量．χ^2_η 的均值和方差分别是 $E\{\chi^2_\eta\}=\eta$ 和 $\mathrm{var}\{\chi^2_\eta\}=2\eta$.

令 $Q_\eta(p)$ 表示 χ_η^2 的概率密度函数的 $p\times100\%$ 百分点：

$$P[\chi_\eta^2 \leqslant Q_\eta(p)] = p.$$

在表 263 中列出了对于自由度 η 从 1 到 7.5(步长为 0.5)、要求置信度分别为 90%、95% 和 99% 的百分点，这些值将在 8.4 节中用到．当 η 较大的时候，$Q_\eta(p)$ 的值可以用下面的近似式来计算(Chambers 等，1983)：

$$Q_\eta(p) \approx \tilde{Q}_\eta(p) \equiv \eta\left(1 - \frac{2}{9\eta} + \Phi^{-1}(p)\left(\frac{2}{9\eta}\right)^{1/2}\right)^3, \tag{264}$$

其中 $\Phi^{-1}(p)$ 是表中最后一行给出的标准高斯分布的 $p\times100\%$ 分点．对所有的 $\eta\geqslant8$，在表中所列 p 的所有六个值的相对误差 $|\tilde{Q}_\eta(p)-Q_\eta(p)|/Q_\eta(p)$ 都小于 0.05．对任意的 η 和 p，高精度的计算 $Q_\eta(p)$ 的算法由 Best and Roberts(1975) 给出，它也是上表中数据的来源．

χ^2 分布经常作为高斯随机变量的平方的非负线性组合的逼近部分的均值．将在 8.4 节式(313a)邻近的讨论中给出如何阐述这种逼近的一个例子．

7.3 一种贝叶斯观点

也许总结一些后面(第 10 章)所需要的贝叶斯(Bayes)想法和术语的最简单的方法是通过下面的途径．假定我们观察到随机变量 X 有依赖于 θ 的概率密度函数；即 θ 为概率密度函数的一个参变量．例如，假定 θ 是一个方差为 1 的高斯随机变量的均值．从一个频率论者或经典统计学的观点来说，这可以表示为(参见式(257a))

$$X \stackrel{\mathrm{d}}{=} \mathcal{N}(\theta,1),$$

其中 θ 为固定的未知量．然而，从贝叶斯的角度来看，在我们对 θ 的值有一定概念(即先验信息)的情况下，θ 被认为是随机的，这种模型象征性地表示为

$$X \mid \theta \stackrel{\mathrm{d}}{=} \mathcal{N}(\theta,1).$$

因为此时 θ 被看作是一个随机变量，它自然就有一个概率密度函数，称为先验概率密度函数，例如

$$\theta \stackrel{\mathrm{d}}{=} \mathcal{N}(0,\sigma_\theta^2).$$

这里 σ_θ^2 是先验分布的一个参数——贝叶斯文中的一个超参数．(这些观点很自然地扩展到多变量随机向量 $\boldsymbol{X}$，其中 $\boldsymbol{X}$ 的联合分布包括一个超参数向量．)

• *方差混合*

在贝叶斯观点下，拉普拉斯(Laplace)概率密度函数作为高斯分布的"方差混合"而出现，这个方差具有一个(单边)指数概率密度函数(参见 10.4 节和 10.5 节)．令

$$X \mid \sigma^2 \stackrel{\mathrm{d}}{=} \mathcal{N}(0,\sigma^2),$$

并且假定随机变量 σ^2 有指数概率密度函数

$$f_{\sigma^2}(\sigma_0^2;v)=v\mathrm{e}^{-\sigma_0^2 v},\quad v>0,$$

这里 v 是一个超参数. 那么高斯型分布的混合方差可以通过对所有的 σ_0^2 的值积分而得到，并由下式给出：

$$\int_0^{\infty} f_{X|\sigma^2=\sigma_0^2}(x)f_{\sigma^2}(\sigma_0^2;v)\mathrm{d}\sigma_0^2=\frac{1}{\sqrt{2/v}}\mathrm{e}^{-(2v)^{1/2}|x|} \tag{265a}$$

(Teichroew, 1957). 这个积分精确地给出了当 $\mu=0$ 时式(257c)的拉普拉斯概率密度函数或双边指数概率密度函数. 如果高斯分布是以 μ 而不是以 0 为中心对称，就恰好得到式(257c)的准确值. 从式(260)，我们可以将上述积分解释为随机变量 X 和 σ^2 的联合概率密度函数，因此这个积分给出了 X 的边际概率密度函数，即拉普拉斯概率密度函数.

• 贝叶斯风险和贝叶斯规则

基于结果 $X=x$，假定当 $\theta=\theta_0$ 时，决策 $d(x)$ 导致了损失 $\ell(d(x),\theta_0)$. 如果我们对 X 的所有可能的结果都取期望(θ 取固定的 θ_0)，就得到一个风险函数

$$r_d(\theta_0)=E_{X|\theta=\theta_0}\{\ell(d(X),\theta_0)\}=\int\ell(d(x),\theta_0)f_{X|\theta=\theta_0}(x)\mathrm{d}x.$$

与 $d(\cdot)$ 有关的贝叶斯风险 R_d 是 $r_d(\theta_0)$ 的期望(θ_0 取所有可能的值)；也就是说，风险函数是 θ 的先验概率密度函数 $f_\theta(\cdot)$ 及其加权积分.

练习[265] 证明

$$R_d\equiv E_\theta\{r_d(\theta)\}=\int\left(\int\ell(d(x),\theta_0)f_{\theta|X=x}(\theta_0)\mathrm{d}\theta_0\right)f_X(x)\mathrm{d}x. \quad\triangleleft$$

圆括号中的积分称为是后验期望损失——这个 x 的函数被概率密度函数 $f_X(\cdot)$ 加权而获得了贝叶斯风险. 根据定义，贝叶斯规则就是选择合适的 $d(x)$ 以最小化贝叶斯风险. 当 $d(x)$ 的选择比较恰当，以至于对所有的 x 后验期望损失都取得极小值的时候，风险最小.

贝叶斯规则可以用来寻找 θ 的点估计，即从数据中导出的单个数值. 贝叶斯决策规则在这种情况下称为是贝叶斯估计. 让我们考虑平方误差损失的贝叶斯估计；即 $\ell(d(x),\theta)=(d(x)-\theta)^2$. 此时，后验期望损失是均方误差：

$$\int_{-\infty}^{\infty}\ell(d(x),\theta_0)f_{\theta|X=x}(\theta_0)\mathrm{d}\theta_0=\int_{-\infty}^{\infty}(d(x)-\theta_0)^2 f_{\theta|X=x}(\theta_0)\mathrm{d}\theta_0.$$

练习[261]告诉我们当 $d(x)$ 等于具有概率密度函数 $f_{\theta|X=x}(\cdot)$ 的随机变量的期望值的时候，上式右面的积分取得最小值. 所以，贝叶斯规则(记作 $B_2(\cdot)$)为

$$B_2(x)\equiv E\{\theta\mid X=x\}. \tag{265b}$$

在平方误差损失的情况下，θ的贝叶斯估计是θ的后验均值. 要注意的是，如果我们将X和θ看作是二元随机变量，类似于式(261a)中的X_0和X_1，则$B_2(x)$是给定$X=x$的条件下θ的条件均值.

作为第二个例子，让我们考虑绝对误差损失的贝叶斯估计量，即$\ell(d(x),\theta)=|d(x)-\theta|$，则后验期望损失是平均绝对误差：

$$\int_{-\infty}^{\infty}\ell(d(x),\theta_0)f_{\theta|X=x}(\theta_0)\mathrm{d}\theta_0=\int_{-\infty}^{\infty}|d(x)-\theta_0|f_{\theta|X=x}(\theta_0)\mathrm{d}\theta_0.$$

极小化上面式子的贝叶斯规则$B_1(\cdot)$是

$$\int_{-\infty}^{B_1(x)}f_{\theta|X=x}(\theta_0)\mathrm{d}\theta_0=0.5 \tag{266}$$

的解. 所以这个例子中的贝叶斯估计量就是θ的后验中值. 同理，如果我们将X和θ看作是二元随机变量，类似于式(261b)中的X_0和X_1，则$B_1(x)$是给定$X=x$的条件下θ的条件中值.

贝叶斯统计有很多很好的教科书. 常见的有 Box and Tiao(1973)，Lee(1989)和 Gelman 等(1995)的书.

7.4 平稳随机过程

现在令$\{X_t: t=\cdots, -1, 0, 1, \cdots\}$是一个离散参数的实值随机过程，按照定义它是以整数为下标的一列随机变量(当独立参变量取值范围为整个实数轴的时候，我们称这个随机过程是连续参数的). 如$\{X_t\}$的过程可以作为一些物理现象一系列观测的随机模型. 我们假设这些观测以步长为Δt记录了下来，假设这个步长具有一定的物理意义，比如说是秒或者年.

• 平稳性和自协方差序列

过程$\{X_t\}$是(二阶)平稳的，如果它满足如下性质：

[1]对所有的整数t，都有$E\{X_t\}=\mu_X$；也就是说，过程$\{X_t\}$的第t个分量X_t的期望值是一个有限的常数μ_X，它不依赖于t.

[2]对所有的整数t和τ，都有$\mathrm{cov}\{X_t, X_{t+\tau}\}=s_{X,\tau}$；也就是说，过程$\{X_t\}$的第$t$个和第$t+\tau$个分量的协方差是一个有限的常数$s_{X,\tau}$，它只依赖于两个分量的指标$t$与指标$t+\tau$之间的间隔$\tau$.

性质[2]的一个重要的推论是：平稳过程$\{X_t\}$的第t个分量X_t的方差是一个与t无关的常数，因为

$$\mathrm{var}\{X_t\}=\mathrm{cov}\{X_t,X_t\}=E\{(X_t-\mu_X)^2\}=s_{X,0}.$$

序列$\{s_{X,\tau}: \tau=\cdots, -1, 0, 1, \cdots\}$称为是自协方差序列(ACVS). 这个序列关于$\tau=0$是对称的，即对所有的$\tau$都有$s_{X,-\tau}=s_{X,\tau}$. 如果我们令$\rho_{X,\tau}\equiv s_{X,\tau}/s_{X,0}$，就得到了一个相关系数序列$\{\rho_{X,\tau}\}$，这就是自相关序列(ACS).

• 谱密度函数

在物理学中作为模型、感兴趣的许多实值平稳过程，包含在$\{X_t\}$的自协方差序列中的关于X_t的信息可以利用谱密度函数$S_X(\cdot)$(SDF)重新表示出来，它还常常称为是功率谱(有时候简称为谱)．如果自协方差序列是平方可和的，即

$$\sum_{\tau=-\infty}^{\infty} s_{X,\tau}^2 < \infty, \tag{267a}$$

我们就有如下关系：

$$S_X(f) = \Delta t \sum_{\tau=-\infty}^{\infty} s_{X,\tau} e^{-i2\pi f\tau\Delta t}, \quad |f| \leqslant f_N \equiv \frac{1}{2\Delta t}, \tag{267b}$$

其中f_N称作奈奎斯特频率，并且单位是每秒几个周期(自协方差序列的平方可和是谱密度存在的充分不必要条件)．当Δt取单位步长的时候，将上式与(35b)相比，可以证明$S_X(\cdot)$是$\{s_{X,\tau}\}$的傅里叶变换，甚至当$\Delta t \neq 1$的时候也是如此．谱密度函数是频率f的偶函数($S_X(-f)=S_X(f)$)并且满足

$$\int_{-f_N}^{f_N} S_X(f) e^{i2\pi f\tau\Delta t} df = s_{X,\tau}, \quad \tau = \cdots, -1, 0, 1, \cdots \tag{267c}$$

(如果$\{s_{X,\tau}\}$是平方可和的并且Δt是1，这个关系其实就是傅里叶逆变换．但是，即使是谱密度函数存在但是无法用(267b)中的自协方差序列表示，上面的关系仍然成立)．特别地，当$\tau=0$的时候，我们有一个基本的结果：

$$\int_{-f_N}^{f_N} S_X(f) df = s_{X,0} = \operatorname{var}\{X_t\}; \tag{267d}$$

即按照频率，$S_X(\cdot)$分解了过程的方差．数量$S_X(f)\Delta f$可以看作是包含在小区间Δf中的频率对过程的方差所做的贡献．实值函数$S_X(\cdot)$是抽样步长为Δt的实值平稳过程的谱密度函数的要求十分简单．仅仅需要对所有f，均满足$S_X(f) \geqslant 0$，$S_X(f)=S_X(-f)$，并且$0 \leqslant \int_{-f_N}^{f_N} S_X(f) df < \infty$，这个积分就是过程的方差(积分为零的例子是没有实际意义的)．

• 线性滤波

为了方便，现在我们令$\Delta t=1$，然后考虑一下对平稳过程$\{X_t\}$滤波的效果．假设$\{a_t: t=0, 1, \cdots, M-1\}$是一个宽度为$M$的线性滤波器，其传递函数定义为

$$A(f) \equiv \sum_{t=0}^{M-1} a_t e^{-i2\pi ft}.$$

如果我们构造了随机过程

$$U_t \equiv \sum_{l=0}^{M-1} a_l X_{t-l}, \tag{268a}$$

那么，由此得到$\{U_t\}$是一个平稳过程，它的谱密度函数为

$$S_U(f) \equiv \mathcal{A}(f) S_X(f), \tag{268b}$$

其中，$\mathcal{A}(f) \equiv |A(f)|^2$ 为对应 $A(\cdot)$的平方增益函数. 因为谱密度函数的积分总是等于过程的方差，我们可以看到

$$\mathrm{var}\{U_t\} = \int_{-1/2}^{1/2} S_U(f)\,\mathrm{d}f = \int_{-1/2}^{1/2} \mathcal{A}(f) S_X(f)\,\mathrm{d}f. \tag{268c}$$

式(268b)是4.2节中我们讨论结果的随机形式：如果我们使用一个平方增益函数为$\mathcal{H}(\cdot)$的滤波器，去对一个有限序列$\{X_t: t=0, 1, \cdots, N-1\}$(序列变量是非随机的其离散傅里叶变换为$\mathcal{X}_k$：$k=0, \cdots, N-1$，能量谱为$|\mathcal{X}_k|^2/N$)循环滤波，则滤波结果的能量谱为$\left\{\mathcal{H}\left(\frac{k}{N}\right)|\mathcal{X}_k|^2/N\right\}$. 类似地，式(268c)是滤波的输出序列的能量为$\sum_{k=0}^{N-1} \mathcal{H}\left(\frac{k}{N}\right)|\mathcal{X}_k|^2/N$的随机形式.

只要式(268c)中的积分 $\mathrm{var}\{U_t\}$是一个有限值，用一个有限长滤波器滤波一个平稳随机过程的结果再次产生了一个平稳过程.

• *白噪声*

一个特别感兴趣的平稳随机过程是白噪声过程$\{\varepsilon_t\}$(又称为完全随机过程). 根据定义，$\{\varepsilon_t\}$是两两不相关的随机变量组成的序列，它们具有同样的均值(常常令它为零)和同样的方差 σ_ε^2. 因为不相关意味着对所有的 t 和 $\tau \neq 0$，都有 $\mathrm{cov}\{\varepsilon_t, \varepsilon_{t+\tau}\}=0$，我们可以看到，对于$\{\varepsilon_t\}$的自协方差序列是

$$s_{\varepsilon,\tau} = \mathrm{cov}\{\varepsilon_t, \varepsilon_{t+\tau}\} = \begin{cases} \sigma_\varepsilon^2, & \tau = 0; \\ 0, & \text{其他}. \end{cases} \tag{268d}$$

因为这个自协方差序列显然是平方可和的，从式(267b)得到，对于 $|f| \leqslant f_{\mathcal{N}}$，$\{\varepsilon_t\}$具有由 $S_\varepsilon(f)=\sigma_\varepsilon^2 \Delta t$ 给出的谱密度函数.

• *自回归过程*

随机过程$\{X_t\}$称为是 p 阶自回归(AR)，如果

$$X_t = \sum_{n=1}^{p} \phi_{p,n} X_{t-n} + \varepsilon_t, \tag{268e}$$

其中，$\{\varepsilon_t\}$是均值为零、方差为 σ_ε^2 的白噪声过程. 我们常常称之为 AR(p)过程. 如果自回归系数 $\phi_{p,n}$满足一定的约束条件(例如，Priestley，1981，3.5.4节)，则$\{X_t\}$是一个平稳过程，并且其谱密度为

$$S_X(f) = \frac{\sigma_\varepsilon^2 \Delta t}{\left|1 - \sum_{n=1}^{p} \phi_{p,n} \mathrm{e}^{-\mathrm{i}2\pi f n \Delta t}\right|^2}, \quad |f| \leqslant f_{\mathcal{N}}. \tag{268f}$$

我们将在7.9节讨论平稳自回归过程的情况.

7.5 谱密度估计

谱密度函数的估计是很多教科书的主题(例如，参见 Blackman and Tukey，1958；Jenkins and Watts，1968；Koopmans，1974；Bloomfield，1976；Priestley，1981；Marple，1987；Kay，1988 及 Percival and Walden，1993). 这里我们简要介绍一下这个主题，作为第 8 章到第 10 章的背景. 我们首先总结周期图的性质，从它是谱密度函数的不相容估计量的意义上来说，它是一个“粗”估计量(也就是说，它的方差不会随着样本容量的增加而减少). 然后，我们将简要介绍一下著名的利用沿着频率平均周期图(即平滑或者滤波)构造相容估计量的方法. 所得到的估计量的特点是，它具有光滑的带宽，这可以用频率的有效范围来度量，在这个范围上周期图是平均得到的. 这个估计量的缺点是，即使谱密度函数不同部分的频谱形状可能表现出不同的急剧或者缓慢的变化，但在整个谱密度函数上它具有固定的光滑带宽. 然后我们讨论第二个“未处理”的谱密度函数的估计量，也就是多锥形估计量. 这个估计量将在 10.7 节起到十分重要的作用，在那里我们将讨论令人感兴趣的基于小波构造更光滑可变宽度的方法.

• *周期图*

给定一个时间序列，它可以看作是一个均值为零的实值平稳随机过程$\{X_t\}$的部分 X_0，X_1，…，X_{N-1} 的实现，它的自协方差是$\{s_{X,\tau}\}$，并且其谱密度函数 $S_X(\cdot)$满足式(267b)，我们利用所谓的有偏的自协方差序列估计量就可以对 $|\tau|=0$，1，…，$N-1$ 估计出 $s_{X,\tau}$，但是无法估计 $|\tau|\geqslant N$ 的 $s_{X,\tau}$，有偏的自协方差估计量是

$$\hat{s}^{(p)}_{X,\tau}=\frac{1}{N}\sum_{t=0}^{N-|\tau|-1}X_tX_{t+|\tau|}$$

(这里明显地使用了过程的均值为零的假设). 自然地，对于 $|\tau|\leqslant N-1$，我们可以将式(276b)中的 $s_{X,\tau}$用 $\hat{s}^{(p)}_{X,\tau}$代替，并且将求和指标限制在$\pm(N-1)$，这也等价于定义 $\hat{s}^{(p)}_{X,\tau}=0$，$|\tau|\geqslant N$. 现在，通过两层求和中的变量变换，我们就可以得到

$$\begin{aligned}\Delta t\sum_{\tau=-(N-1)}^{N-1}\hat{s}^{(p)}_{X,\tau}\mathrm{e}^{-\mathrm{i}2\pi f\tau\Delta t}&=\frac{\Delta t}{N}\sum_{\tau=-(N-1)}^{N-1}\sum_{t=0}^{N-|\tau|-1}X_tX_{t+|\tau|}\mathrm{e}^{-\mathrm{i}2\pi f\tau\Delta t}\\&=\frac{\Delta t}{N}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1}X_jX_k\mathrm{e}^{-\mathrm{i}2\pi f(k-j)\Delta t}\\&=\frac{\Delta t}{N}\left|\sum_{t=0}^{N-1}X_t\mathrm{e}^{-\mathrm{i}2\pi ft\Delta t}\right|^2\equiv\hat{S}^{(p)}_X(f),\end{aligned}\tag{269a}$$

上面定义的函数$\hat{S}^{(p)}_X(\cdot)$就称为是*周期图*(尽管它是频率的函数而且不是周期的). 和 $S_X(\cdot)$一样，它定义在区间$[-f_{\mathcal{N}},f_{\mathcal{N}}]$上. 可以证明(参见 Brillinger，1981，5.2 节)，在一定的高阶矩有限的条件下，当 $N\to\infty$的时候，渐进地有

$$\hat{S}^{(p)}_X(f)\stackrel{\mathrm{d}}{=}\begin{cases}S_X(f)\chi^2_2/2, & 0<f<f_{\mathcal{N}};\\ S_X(f)\chi^2_1, & f=0\text{ 或 }f_{\mathcal{N}},\end{cases}\tag{269b}$$

其中，如7.1中所说，χ_η^2表示具有自由度为η的χ^2分布的随机变量．注意，因为$E\{\chi_\eta^2\}=\eta$，由上面得到，对于大的N有$E\{\hat{S}_X^{(p)}(f)\}\approx S_X(f)$；也就是说，渐进地周期图变成了谱密度函数的无偏估计，因为它的期望与我们要估计的量是一致的．进而，对于$0\leqslant f'<f\leqslant f_N$，$\hat{S}_X^{(p)}(f)$和$\hat{S}_X^{(p)}(f')$是渐进独立的．利用结果$\mathrm{var}\{\chi_\eta^2\}=2\eta$，我们有：当$N\to\infty$时，渐进地

$$\mathrm{var}\{\hat{S}_X^{(p)}(f)\}=\begin{cases}S_X^2(f), & 0<f<f_N;\\ 2S_X^2(f), & f=0\text{ 或者 }f=f_N;\end{cases}\tag{270a}$$

$$\mathrm{cov}\{\hat{S}_X^{(p)}(f),\hat{S}_X^{(p)}(f')\}=0,\quad 0\leqslant f'<f\leqslant f_N,$$

(对于有限的样本数N，如果特定的约束满足，这些渐进的结果是十分有用的逼近，参见Percival and Walden，1993，p. 232)．

周期图看来具有一个明显的主要缺点和一个明显的主要优点．因为，一般情况下$S_X(f)>0$，所以(270a)证明当$N\to\infty$的时候，$\hat{S}_X^{(p)}(f)$是不会减少到零的，也就是说$\hat{S}_X^{(p)}(f)$变得任意接近它的期望的概率是零．正因为如此，$\hat{S}_X^{(p)}(f)$是$S_X(f)$的不相容估计量．另一方面，周期图中任意两项的渐进独立性，使得理解沿着频率平滑周期图构造的谱密度函数估计量的采样性质成为可能．

周期图的第二个——并不是最明显的——优点是，我们能很容易地将随机变量$\hat{S}_X^{(p)}(f)$变换为一个新的方差不依赖于未知的谱密度函数的随机变量(式(270a)表明周期图本身没有这个性质)．借助于下面的讨论，从式(270a)可以导出一个结论：对于周期图的"稳定方差"变换恰好就是这个算法．因为当$0<f<f_N$的时候，$\hat{S}_X^{(p)}(f)$和$S_X(f)\chi_2^2/2$有同样的分布，由此得出$\log(\hat{S}_X^{(p)}(f))$和$\log(S_X(f)\chi_2^2/2)=\log(S_X(f)/2)+\log(\chi_2^2)$有同样的分布．因为一个随机变量加上一个常数的方差等于这个随机变量的方差本身，所以$\mathrm{var}\{\log(\hat{S}_X^{(p)}(f))\}=\mathrm{var}\{\log(\chi_2^2)\}$．基于此，Bartlett and Kendall(1946)证明了，对于$0<f<f_N$，有

$$\mathrm{var}\{\log(\hat{S}_X^{(p)}(f))\}=\frac{\pi^2}{6},\quad E\{\log(\hat{S}_X^{(p)}(f))\}=\log(S_X(f))-\gamma,$$

其中$\gamma\doteq0.577\,21$是欧拉(Euler)常数．因此，对数周期图是"稳定方差过程"．对于$0<f<f_N$，随机变量

$$\varepsilon(f)\equiv\log\left(\frac{\hat{S}_X^{(p)}(f)}{S_X(f)}\right)+\gamma\tag{270b}$$

具有零均值且方差为$\sigma_\varepsilon^2=\pi^2/6$；进而，

$$\varepsilon(f)\stackrel{\mathrm{d}}{=}\log(\chi_2^2)+\gamma-\log(2)\tag{270c}$$

(练习[7.5]中的验证). 如果令

$$Y^{(p)}(f) \equiv \log(\hat{S}_X^{(p)}(f)) + \gamma,$$

那么对于 $0<f<f_N$，有

$$Y^{(p)}(f) = \log(S_X(f)) + \varepsilon(f). \tag{271a}$$

所以，对数周期图(加上已知的常数 γ)可以写成一个“信号”(准确地说，是谱密度函数的对数)加上具有零均值且已知的方差为 $\sigma_\varepsilon^2=\pi^2/6$ 的非高斯噪声.

• 传统的周期图平滑

谱密度函数的相容估计量可以通过周期图的纵坐标在一定方式下的组合产生. 例如，这个问题的传统的方法是在傅里叶频率域 $f_j=j/(N\Delta t)$ 计算周期图的平滑(光滑). 所得到的 $\lfloor N/2 \rfloor+1$ 个随机变量

$$\hat{S}_X^{(p)}(f_0), \hat{S}_X^{(p)}(f_1), \cdots, \hat{S}_X^{(p)}(f_{\lfloor N/2 \rfloor})$$

当 N 充分大的时候，都是近似两两独立的；即

$$\operatorname{cov}\{\hat{S}_X^{(p)}(f_j), \hat{S}_X^{(p)}(f_k)\} \approx 0, \quad j \neq k \text{ 且 } 0 \leqslant j,k \leqslant \lfloor N/2 \rfloor.$$

因此，

$$\hat{S}_X^{(p)}(f_j) \stackrel{d}{=} S_X(f_j)U_j, \quad j = 0,1,\cdots,\lfloor N/2 \rfloor, \tag{271b}$$

其中 U_j，$j=1$，$\cdots$,$\lfloor (N-1)/2 \rfloor$，是渐近的独立同分布 $\chi_2^2/2$(作为对比，U_0 和 $U_{N/2}$——如果 N 为偶数——的分布都是 χ_1^2). 另一方面，如果我们用式(270b)定义 $\varepsilon(f_j)$，那么 $\{\varepsilon(f_j): 0<f_j<f_N\}$ 组成了一族近似地具有零均值和方差为 $\pi^2/6$ 的独立同分布随机变量(注意 U_j 和 $\varepsilon(f_j)$ 都不是高斯分布).

假设 N 充分大，使周期图 $\hat{S}_X^{(p)}(\cdot)$ 是 $S_X(\cdot)$ 的无偏估计，并且在傅里叶频率 f_j 中是两两不相关的. 如果 $S_X(\cdot)$ 在 f_j 邻域内的变化很缓慢，那么对于某个整数 $M>0$，有

$$S_X(f_{j-M}) \approx \cdots \approx S_X(f_j) \approx \cdots \approx S_X(f_{j+M}),$$

因此

$$\hat{S}_X^{(p)}(f_{j-M}), \cdots, \hat{S}_X^{(p)}(f_j), \cdots, \hat{S}_X^{(p)}(f_{j+M})$$

是由 $2M+1$ 个数组成的对同一个量 $S_X(f_j)$ 的一族无偏的和不相关的估计量. 因此我们能够取它们的平均以产生估计量

$$\bar{S}_X(f_j) \equiv \frac{1}{2M+1}\sum_{l=-M}^{M} \hat{S}_X^{(p)}(f_{j-l}). \tag{271c}$$

在我们的假设下，利用式(270a)并且为了简单起见假设 $f_{j-M}>0$ 和 $f_{j+M}<f_N$，则有

$$E\{\bar{S}_X(f_j)\} \approx S_X(f_j)$$

和

$$\operatorname{var}\{\bar{S}_X(f_j)\} \approx \frac{S_X^2(f_j)}{2M+1} \approx \frac{\operatorname{var}\{\hat{S}_X^{(p)}(f_j)\}}{2M+1}.$$

表 272

	$c=1$	$c=2$	$c=3$	$c=4$
$r=0$	2.521 628 1	−4.771 535 9	7.919 991 5	−11.976 921 1
$r=1$	16.077 882 8	−20.634 334 6	25.053 152 1	−28.873 813 6
$r=2$	31.804 626 5	−34.007 137 3	34.770 027 2	−34.315 132 1
$r=3$	32.786 109 9	−30.286 123 3	26.710 935 6	−22.883 831 0
$r=4$	18.743 209 8	−14.571 768 8	10.717 774 4	−7.532 219 4
$r=5$	4.722 631 9	−2.680 792 3	1.339 130 6	−0.516 712 5

注：AR(24)过程(Gao，1997)的系数$\{\phi_{24,n}: n=1, \cdots, 24\}$．第$r$行第$c$列的系数是$\phi_{24,4r+c}$．这些系数可以在本书的网页上找到.

现在，如果我们考虑增加样本容量N和下标j保证$j/(N\Delta t)=f_j$是一个常数，那么可以令M变大，并且我们可以断言，$\mathrm{var}\{\bar{S}_X(f_j)\}$可以任意小，于是$\bar{S}_X(f_j)$是$S_X(f_j)$的相容估计量．上述结论的一个明显的推论就是在式(271c)的求和中使用非均匀的权重(这种和其他的一些改进可以参考 Cleveland and Parzen，1975；Bloomfield，1976；Walden，1990 及 Percival and Walden，1993.

Wahba(1980)使用了不同的方法，它研究了利用光滑样条对对数周期图的平滑．这种方法使用了式(271a)中的信号加噪声模型，然而，回忆一下噪声具有$\log(\chi_2^2)$分布，所以这个模型很复杂.

对数周期图的平滑和样条平滑都有如下特征：都有平滑的带宽，它度量了在多大的带宽上潜在的谱密度函数估计进行平滑处理而构造相容估计量．例如，均匀周期图平滑将相邻的$2M+1$个数进行了平均，而每一个数本身也是$1/(N\Delta t)$上的平均，所以带宽是$(2M+1)/(N\Delta t)$．对特别的谱密度函数，常常以一种满意的方式确定带宽，通过平滑引起的偏差和满足要求的方差的折衷来确定它的带宽．尽管如此，在整个谱密度函数上的带宽是固定的，并没有随着谱密度函数的局部特征而有任何形式的改变，例如，急剧的尖峰值或者凹谷需要窄带宽平滑以避免平滑引起的优势偏差，而谱密度函数变换缓慢的区域使用宽带宽的平滑就可以大幅度减少方差．正如我们将在10.6节和10.7节中看到的那样，一种产生自适应平滑谱密度函数估计量的方法，就是使用小波将谱密度函数中各种不同分辨率(尺度)中的信息融合起来.

• *多锥谱密度函数估计*

如我们能够由式(269b)推出的，统计理论认为，对于大的样本容量N的平稳随机过程，周期图是真正的谱密度函数的近似的无偏估计量．然而，重要的是意识到，根据所谓的漏泄现象对于有限的样本容量的过程，这种逼近是十分无力的．图273通过比较一个长度$N=2048$的时间序列的周期图(细的锯齿状线)和真正的谱密度函数(粗的平滑线)演示了漏泄效应，这个时间序列是$\sigma_\epsilon^2=1$，系数$\{\phi_{24,n}\}$为表272所列的AR(24)过程的一个实现(这个过程在第10章和Gao(1997)

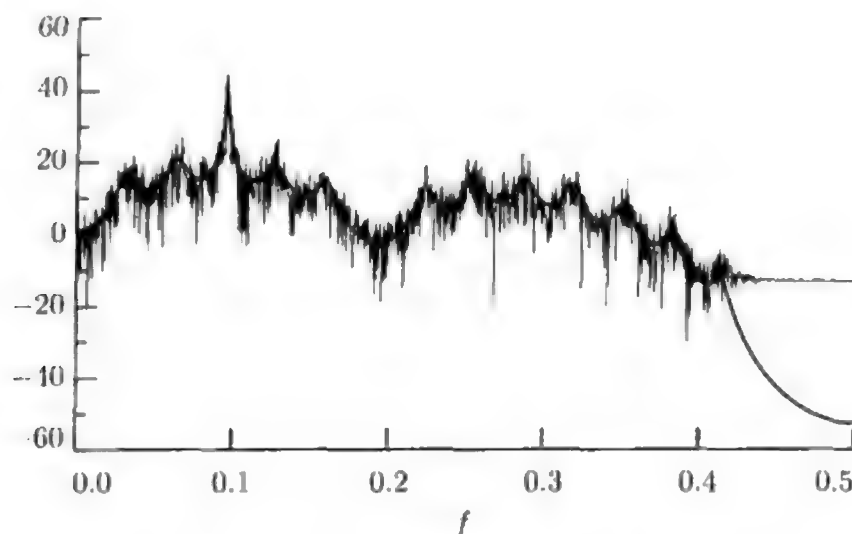


图 273　长度 $N=2\,048$ 的时间序列，AR(24)过程的实现(参见表 272 中定义这个过程的系数)的周期图(锯齿状细线)和真正的谱密度函数(光滑的粗线)．周期图和真正的谱密度函数都绘制在分贝尺度的图上．周期图的高频中的泄漏是显然的，这里的偏差变得大约有 40dB(即 4 数量级)

中也用到了．生成自回归过程的实现的细节，参见 7.9 节)．周期图和真正的谱密度函数都绘制在分贝(dB)尺度上；即绘制的是 $10\cdot\log_{10}(\hat{S}_X^{(p)}(f))$ 作为 f 的函数图像．泄漏的表现是，在高频的地方周期图和真正的谱密度函数的偏差很大．一般地说，对谱密度函数分布占据较大范围的随机过程(也就是说谱密度函数最大值与最小值的比值大于 40 或 50 分贝，比如说这个 AR(24)过程的比值接近 90 分贝)来说，周期图的泄漏现象是十分危险的．

一个公认的减小由于周期图漏泄导致的偏差的过程是在计算频谱估计量前对时间序列进行数据锥(窗)处理，这就产生了所谓的直接频谱估计量．直接频谱估计量比周期图具有更好的小样本偏差小属性；但是，当接着对频率域进行平滑的时候，渐近方差要比周期图平滑后要大．多锥频谱估计的基本思想是，用计算少量的 K 个不同数据锥分别计算直接频谱估计然后求 K 个估计量的平均，这样构造的频谱估计的方差比较小(Thomson，1982；Percival and Walden，1993，第 7 章)．如果所有这 K 个锥两两正交，并且恰好阻止了泄漏，根据减少偏差和方差来说，得到的多锥估计就要比周期图好，尤其是频谱的动态范围比较大并且/或者频谱变化快的时候．多锥估计在研究，例如二氧化碳与全球温度的关系(Kuo 等，1990)，血浆波动紊乱(Riedel and Sidorenko，1995)，以及海浪高度(Walden 等，1995)中取得了成功．

和前面一样，令$\{X_t\}$为一个零均值实值平稳过程．令$\{a_{n,t}:\ t=0,\ \cdots,\ N-1\}$，$n=0,\ \cdots,\ K-1$ 表示 K 个不同的数据锥，用于形成多锥谱密度函数估计量．这些锥都是规范正交的，即

$$\sum_{t=0}^{N-1} a_{n,t}a_{l,t} = \begin{cases} 1, & n=l; \\ 0, & \text{其他}. \end{cases}$$

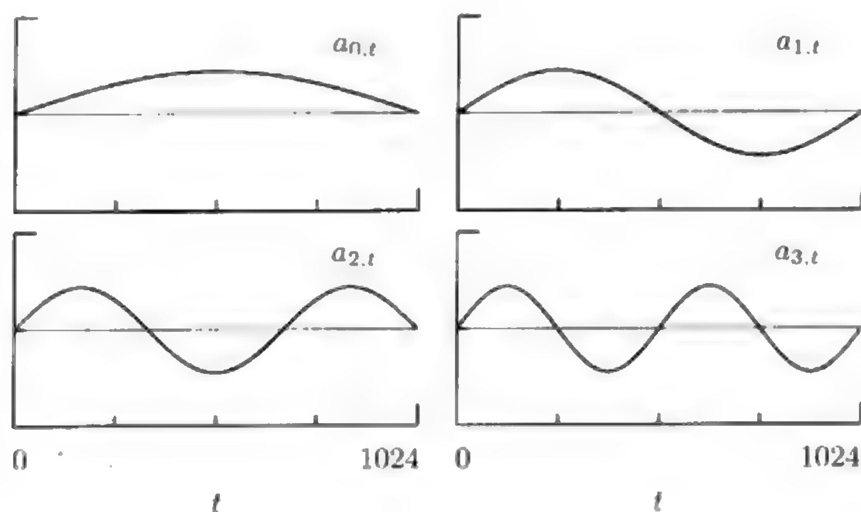


图 274　当 $N=1\ 024$ 时，阶数 $n=0$，1，2 和 3 的正弦锥$\{a_{n,t}\}$

每一个锥的基本任务是产生一个偏差不超过周期图的谱密度函数估计量. 最简单的多锥谱密度函数估计量是 K 个直接频谱估计量(特征谱)的平均，所以取如下形式：

$$\hat{S}_X^{(\mathrm{mt})}(f) \equiv \frac{1}{K}\sum_{n=0}^{K-1} \hat{S}_{X,n}^{(\mathrm{mt})}(f),$$

其中

$$\hat{S}_{X,n}^{(\mathrm{mt})}(f) \equiv \Delta t\left|\sum_{t=0}^{N-1} a_{n,t}X_t \mathrm{e}^{-\mathrm{i}2\pi f t\,\Delta t}\right|^2.$$

求均值运算的基本作用是将可能丢失的但是在某单个锥中使用了的"信息"恢复出来；第二个作用是稍微减少周期图带来的变异性.

容易计算的规范正交锥是正弦锥，第 n 个正弦锥的形式是

$$a_{n,t} = \left(\frac{2}{N+1}\right)^{1/2}\sin\left(\frac{(n+1)\pi(t+1)}{N+1}\right),\quad t=0,\cdots,N-1. \tag{274a}$$

图 274 表示当 $N=1\ 024$，$n=0$，1，2，3 时，这些锥的样子. 这些简单的锥是 Riedel and Sidorenko(1995)引入的，而我们将在 10.7 节中使用. 估计量$\hat{S}_X^{(\mathrm{mt})}(\cdot)$有一个伴随的标准化的带宽$(K+1)/(N+1)$(Riedel and Sidorenko，1995；Walden 等，1995). 如果谱密度函数在这个宽度的带宽上没有急剧的变化，则特征谱是近似无关的，这就产生了一个逼近(对于大的 N 是正确的)

$$\hat{S}_X^{(\mathrm{mt})}(f) \overset{\mathrm{d}}{=} \frac{S_X(f)\chi_{2K}^2}{2K},\quad 0<f<f_{\mathcal{N}}. \tag{274b}$$

如果 K 按照合适的比例随着 N 的增加而增加，则多窗估计量可以做到是相容的. 但是，典型的 K 都是固定的小的数(10 或者更小). 因为 K 是固定的，所以

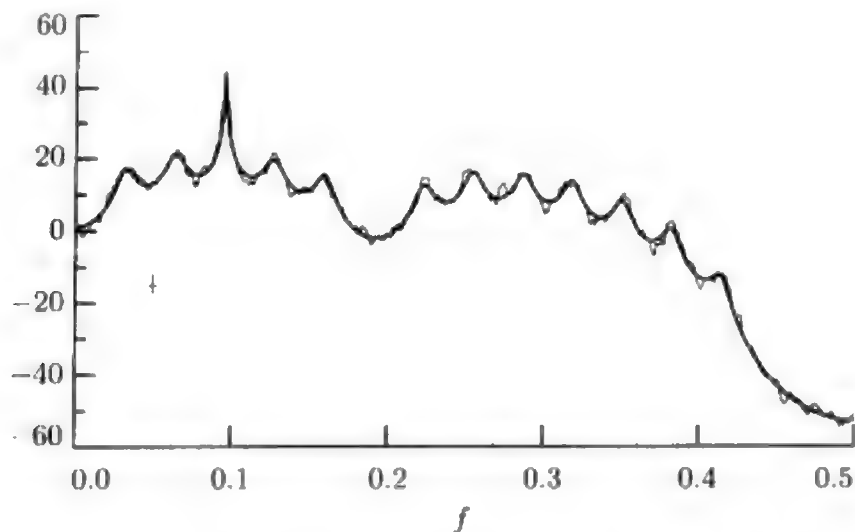


图 275　长度 $N=2\ 048$ 的仿真时间序列 AR(24)(对应的周期图参见图 273)的多锥谱密度函数估计$\hat{S}_X^{(mt)}(\cdot)$(锯齿形细线)和真正的谱密度函数(光滑粗线). 多锥估计是基于 $K=10$ 的正弦锥. $\hat{S}_X^{(mt)}(\cdot)$和真正的谱密度函数绘制在分贝尺度的图上. 图像上左半部分的十字形符号的宽度给出了$\hat{S}_X^{(mt)}(\cdot)$的带宽(即$\frac{K+1}{N+1}\doteq 0.005\ 4$), 在 $10\cdot\log_{10}(S_X(f))$的情况下, 其高度给出了 95%置信区间长度

$\hat{S}_X^{(mt)}(\cdot)$是不相容的谱密度函数估计量; 然而与周期图不同的是, 多锥估计量的偏差较小并且具有 $2K$ 的自由度(K 个特征谱中的任意两个求平均). 多锥估计方法的背景和更多的讨论参见 Thomson(1982), Percival and Walden(1993, 第 7 章).

如图 275 所示, 细线表示同样的 AR(24)时间序列的谱密度函数多锥估计量, 这个时间序列的周期图如图 273 所示(在这两幅图像上, 粗的平滑曲线表示真正的谱密度函数). 这里我们使用 $K=10$ 的正弦锥, 产生了自由度为 $2K=20$ 与带宽为$(K+1)/(N+1)\doteq 0.005\ 4$ 的估计量. 在图的左边, 我们用交叉线的宽度描述带宽, 交叉的高度给出了真正的谱密度函数置信度为 95%的长度, 单位为分贝(这些都基于 χ_{20}^2的分布的百分点, 利用式(264)可以近似地计算出来; 构造和解释这些置信区间的方法, 详情参见 Percival and Walden, 1993, 6.10 节).

如我们注意到的, 对于周期图的一个事实是: 对数变换稳定化多锥谱密度函数估计量的方差. Bartlett and Kendall(1946)关于随机变量 $\log(\chi_{2K}^2)$的性质的结果可以用来证明, 对于 $0<f<f_N$, 有

$$E\{\log(\hat{S}_X^{(mt)}(f))\} = \log(S_X(f)) + \psi(K) - \log(K)$$

和

$$\mathrm{var}\{\log(\hat{S}_X^{(mt)}(f))\} = \psi'(K),$$

其中, $\psi(\cdot)$和 $\psi'(\cdot)$分别表示 Γ_2 函数和 Γ_3 函数:

$$\psi(z) \equiv \frac{\mathrm{d}\log(\Gamma(z))}{\mathrm{d}z}, \qquad \psi'(z) \equiv \frac{\mathrm{d}\psi(z)}{\mathrm{d}z}. \tag{275}$$

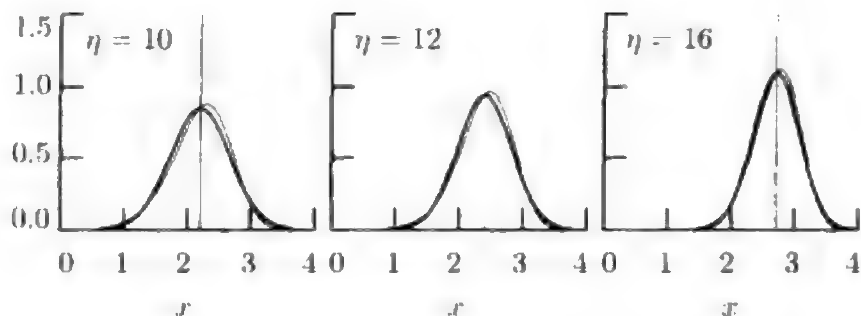


图 276 具有相同均值和方差的 $\log(\chi_\eta^2)$ 随机变量的概率密度函数(细线)和高斯随机变量的概率密度函数(粗线)的比较. 自由度 η 从左到右分别为 10、12 和 16(它们是与多锥谱密度函数估计量 $\hat{S}_X^{(mt)}(\cdot)$, 分别是 $K=5$, 6 和 8 的数据锥, 相关的自由度). 竖直线标出了 $\log(\chi_\eta^2)$ 随机变量的均值——从左到右分别是 $\psi(5)+\log(2)\doteq 2.199$, $\psi(6)+\log(2)\doteq 2.399$, $\psi(8)+\log(2)\doteq 2.709$. 对应的方差的平方根分别是 $\sqrt{\psi'(5)}\doteq 0.470$, $\sqrt{\psi'(6)}\doteq 0.426$, $\sqrt{\psi'(8)}\doteq 0.365$. (练习[7.1]涉及到 $\log(\chi_\eta^2)$ 概率密度函数的导出.)

如图 276 所表明的, 比较概率密度函数 $\log(\chi_{2K}^2)$ 与有着相同期望值和方差的高斯概率密度函数, 我们发现当 $K\geqslant 5$ 的时候, 按照 Bartlett and Kendall(1946)的说法, 它们是很符合的. 对于 $0<f<f_{\mathcal{N}}$, 由此得到随机变量

$$\eta(f)\equiv\log\left(\frac{\hat{S}_X^{(mt)}(f)}{S_X(f)}\right)-\psi(K)+\log(K)$$

是均值为零、方差为 $\psi'(K)$ 的近似高斯分布.

如果对于 $0<f<f_{\mathcal{N}}$, 我们令

$$Y^{(mt)}(f)\equiv\log(\hat{S}_X^{(mt)}(f))-\psi(K)+\log(K),$$

则

$$Y^{(mt)}(f)=\log(S_X(f))+\eta(f). \tag{276}$$

所以, 对数多锥谱密度函数估计量(加上已知常数 $\log(K)-\psi(K)$)可以写成一个信号(真的对数谱密度函数)加上具有零均值和已知方差 $\sigma_\eta^2=\psi'(K)$ 的近似高斯噪声.

如果我们求式(276)在傅里叶频率 $f_k=k/(N\Delta t)$ 上的值, 就会有一个近似的结果, 对于 $0<f_k<f_{\mathcal{N}}$, 有

$$\eta(f_k)\stackrel{d}{=}\mathcal{N}(0,\psi'(K)),$$

但是, 这些随机变量是相关的. 下面我们研究 $\eta(f)$ 在频率域中的协方差结构. 为了符号上的方便, 我们取 $\Delta t=1$(所以 $f_{\mathcal{N}}=1/2$). 对于固定的 f 和 ν, 满足 $0<f<1/2$ 和 $0<f+\nu<1/2$, 让我们定义

$$s_\eta(\nu)\equiv\operatorname{cov}\{\eta(f),\eta(f+\nu)\}=\operatorname{cov}\{\log(V(f)),\log(V(f+\nu))\},$$

其中

$$V(f)\equiv\frac{\hat{S}_X^{(mt)}(f)}{S_X(f)}.$$

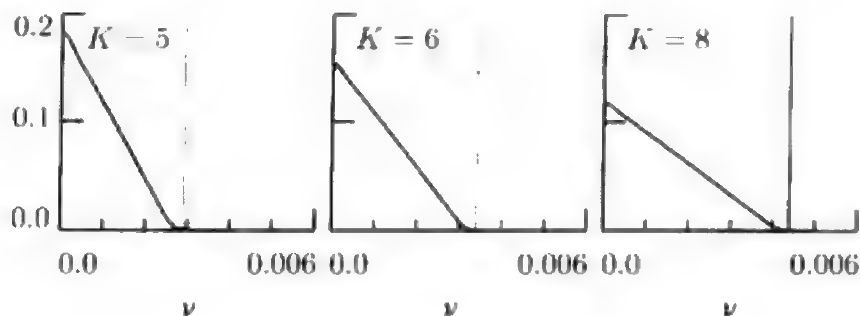


图 277　对于 $N=2\ 048$ 和 $K=5$，6，8 正弦锥，作为 ν 的函数自协方差 $\tilde{s}_{\eta}(\nu)$ 的图像．每个竖线标明了相联系的多锥谱密度函数估计量的带宽 $\frac{K+1}{N+1}$

如果 $\log(V(f))$ 和 $\log(V(f+\nu))$ 是确切地联合严格高斯分布的，我们就有

$$\mathrm{cov}\{\log(V(f)),\log(V(f+\nu))\} = \log(1+\mathrm{cov}\{V(f),V(f+\nu)\})$$

(Granger and Hatanaka，1964，3.7 节)；因为它们是近似地联合高斯分布的，上面我们能合理地近似，所以

$$s_{\eta}(\nu) \approx \log(1+s_V(\nu)),$$

其中

$$s_V(\nu) \equiv \mathrm{cov}\{V(f),V(f+\nu)\}.$$

根据我们的假设，谱密度函数在局部上的变化是慢的，所以对比较小的 ν，$S_X(f)\approx S_X(f+\nu)$ 成立，我们有

$$s_V(\nu) \approx \frac{1}{K^2 S_X^2(f)} \sum_{n=0}^{K-1}\sum_{l=0}^{K-1} \mathrm{cov}\{\hat{S}_{X,n}^{(\mathrm{mt})}(f),\hat{S}_{X,l}^{(\mathrm{mt})}(f+\nu)\}.$$

在对 $S_X(\cdot)$ 做同样假设的条件下，Thomson(1982，p. 1 069)证明了

$$\mathrm{cov}\{\hat{S}_{X,n}^{(\mathrm{mt})}(f),\hat{S}_{X,l}^{(\mathrm{mt})}(f+\nu)\} \approx S_X^2(f)\left|\sum_{t=0}^{N-1} a_{n,t}a_{l,t}\mathrm{e}^{\mathrm{i}2\pi\nu t}\right|^2,$$

这个近似式忽略了频率 f 接近 0 或者 1/2 的时候才显著地依赖于 f 的项．所以，我们有

$$s_V(\nu) \approx \tilde{s}_V(\nu) \equiv \frac{1}{K^2}\sum_{n=0}^{K-1}\sum_{l=0}^{K-1}\left|\sum_{t=0}^{N-1} a_{n,t}a_{l,t}\mathrm{e}^{\mathrm{i}2\pi\nu t}\right|^2.$$

因此，上面的结论给出了 $s_{\eta}(\nu)\approx\tilde{s}_{\eta}(\nu)\equiv\log(1+\tilde{s}_V(\nu))$．

图 277 绘出了当 $N=2\ 048$，$K=5$，6，8 个正弦锥的时候作为 ν 的函数 $\tilde{s}_{\eta}(\nu)$ 的图像．当 $\nu\geqslant(K+1)/(N+1)$ 的时候，$\tilde{s}_{\eta}(\nu)$ 非常小以至于可以忽略，其中 $(K+1)/(N+1)$ 是正弦多锥估计量的标准化带宽，且在上图中用竖直的直线标记出来．自协方差图中的直线形状在其他的 N 和 K 的时候也是如此．回忆 $s_{\eta}(0)=\sigma_{\eta}^2$ 和使用 $(K+1)/(N+1)\approx(K+1)/N$，因此我们能陈述一个非常简单和方便的模型，即

$$s_{\eta}(\nu) = \begin{cases}\sigma_{\eta}^2\left(1-\dfrac{|\nu|N}{K+1}\right), & |\nu|\leqslant(K+1)/N; \\ 0, & \text{其他}.\end{cases} \tag{277}$$

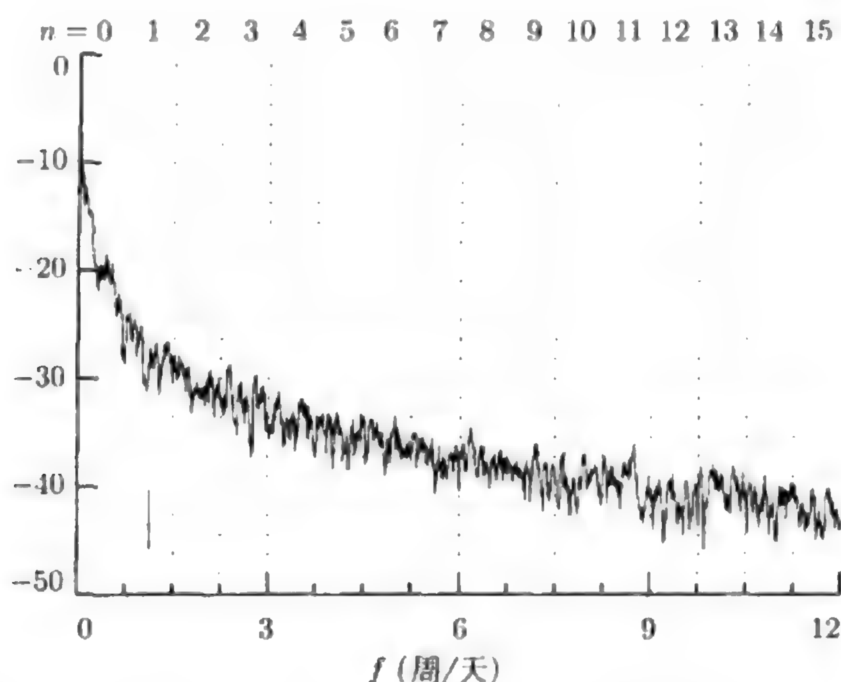


图 278 太阳物理时间序列(这个序列 $N=4\ 096$ 的值绘制在图 222 和图 235)使用 $K=10$ 的正弦锥得到的多锥谱密度函数估计 $\hat{S}_X^{(mt)}(\cdot)$(在分贝尺度上). 竖直排列的点将频率区间[0, 12 周/天]分割为 16 个小区间，如同图 220 中 $j=4$ 的离散小波包变换或者图 236 中极大覆盖离散小波包变换). 图像中左下半部分十字形的宽度给出了 $\hat{S}_X^{(mt)}(\cdot)$的物理带宽(即 $\frac{K+1}{(N+1)\Delta t}\doteq 0.064\ 4$ 周/天——这里 $\Delta t=1/24$ 天)，在 $10\cdot\log_{10}(S_X(f))$的情况下，其高度给出了 95%置信区间长度

7.5 节的评论与扩展

[1]作为谱密度函数的多锥估计量的第二个例子，图 278 展示了 6.2 节中描写过的太阳物理时间序列(这个时间序列具有 $N=4\ 096$ 个数据值，在图 222 和图 235 的底部绘出了这个时间序列的图像)的($K=10$ 个)正弦多锥估计量. 这个估计有 $2K=20$ 的自由度，相联系的带宽为 0.064 4 周/天(通过标准化带宽 $(K+1)/(N+1)$ 和 $1/\Delta t=24$ 的乘积计算得到，其中 $\Delta t=1/24$ 天). 图 278 中左下角的交叉的宽度就表示这个带宽，而交叉的高度就以置信度为 95%给出这个时间序列假定的真正谱密度函数的值. 竖直的点线将频率区间分割成了 16 个小区间，对应着 $j=4$ 层的离散小波变换或者极大重叠离散小波变换(图 236).

对于时间序列的这个或者另外的谱密度函数的估计的局限性在于，我们认为在太阳物理时间序列中观测到的与时间相关的波动与具有时间不相关性的随机过程的实现是一致的(特别地，这就意味着在任何时候观测到的爆发的概率是一样的)；但是，基于 6.2 节所讨论的物理考虑，要使用与时间相关的随机模型. 所以，估计量 $\hat{S}_X^{(mt)}(\cdot)$虽然告诉我们各种频率在时间序列的全局傅里叶分解中的相对重要性. 但是它不可能告诉我们局部的时频关系(离散小波包变换和极大重叠离散

小波包变换可以实现时间局部化，但是它是以增大带宽到0.75周/天为代价的).

7.6 长记忆过程的定义和模型

假设$\{X_t\}$是平稳过程，它的谱密度函数用$S_X(\cdot)$表示(为了方便，我们假设$\Delta t=1$，所以始终有$f_{\mathcal{N}}=1/2$). 我们称$\{X_t\}$是一个平稳长记忆过程，如果存在常数α和C_S，满足$-1<\alpha<0$和$C_S>0$，使得

$$\lim_{f\to 0}\frac{S_X(f)}{C_S\mid f\mid^{\alpha}}=1, \tag{279a}$$

(Beran, 1994, p.42). 换句话说，平稳长记忆过程的谱密度函数$S_X(\cdot)$满足$S_X(f)\approx C_S\mid f\mid^{\alpha}$，当$f$趋于0的时候逼近程度更好. 另外也可以根据$\{X_t\}$的自协方差序列$\{s_{X,\tau}\}$来定义. 我们称$\{X_t\}$是一个平稳长记忆过程，如果存在常数$\beta$和$C_s$，满足$-1<\beta<0$和$C_s>0$，使得

$$\lim_{\tau\to\infty}\frac{s_{X,\tau}}{C_s\tau^{\beta}}=1,$$

其中β与式(279a)中的α有关，$\beta=-\alpha-1$(根据Beran(1994)的定理2.1，这个定义和式(279a)是等价的). 标准的时间序列模型，比如说平稳自回归过程，对于大的τ，其自协方差满足$s_{X,\tau}\approx C\phi^{\tau}$，其中$C\geqslant 0$，$\mid\phi\mid<1$. 对于一个长记忆过程，当$\tau$较大的时候我们有$s_{X,\tau}\approx C_s\tau^{\beta}$. 在这两种情况下都有：随着$\tau\to\infty$，$s_{X,\tau}\to 0$. 但是，对长记忆过程而言衰减到零的速度是非常慢的，也就是说时间序列中距离很远的两个数，仍然有不可忽略的协方差(以更吸引人的术语说，当前观测保留有久远的往事的一些"记忆").

在文献中提出和研究的平稳长记忆过程有若干不同的模型，下面三个是特别重要的.

• 分形高斯噪声

长记忆过程的第一个完整的模型是Mandelbrot和van Ness在1968年提出来的，这就是分形高斯噪声(fractional Gaussian noise, FGN). 根据定义，如果$\{X_t\}$是一个分形高斯噪声，那么它就是一个平稳过程，其自协方差序列是

$$s_{X,\tau}=\frac{\sigma_X^2}{2}(\mid\tau+1\mid^{2H}-2\mid\tau\mid^{2H}+\mid\tau-1\mid^{2H}),\quad \tau=\cdots,-1,0,1,\cdots, \tag{279b}$$

这里$\sigma_X^2=\mathrm{var}\{X_t\}$是一个任意正数，$H$就是满足$0<H<1$的所谓的Hurst(自相似)参数. 分形高斯噪声可以看作是参数为H的分形布朗(Brown)运动(FBM)$\{B_H(t):0\leqslant t<\infty\}$的增量；即

$$X_t=B_H(t+1)-B_H(t),\quad t=0,1,2,\cdots.$$

分形布朗运动的精确定义，例如，在Beran(1994)中给出的，由书中我们注意到$B_H(0)\equiv 0$；对于$t>0$，$B_H(t)$是零均值、方差为$\sigma_X^2t^{2H}$的高斯随机变量；并且当$s\geqslant t\geqslant 0$时$\mathrm{cov}\{B_H(t),B_H(s)-B_H(t)\}=0$. 因此，我们可以建立离散分形布朗运动(DFBM)的样本，而用累积求和$\{X_t\}$，得到

$$B_t\equiv B_H(t)=\sum_{u=0}^{t-1}X_u,\quad t=1,2,\cdots$$

(参见本节的扩展与评论，那里将明确地解释 B_t 和 $B_H(t)$ 中的 t). 对于 $B_H(t)$ 的谱密度函数可以取为

$$S_{B_H(t)}(f)=\frac{\sigma_X^2 C_H}{|f|^{2H+1}},\quad -\infty<f<\infty, \tag{280a}$$

其中 $C_H\equiv\Gamma(2H+1)\sin(\pi H)/(2\pi)^{2H+1}$(参见 Mandelbrot and van Ness (1968)；Flandrin(1989)；Masry(1993))．所以，$S_{B(t)}(f)\propto|f|^{-1-2H}$. 从滤波和混合的考虑出发(见 Percival and Walden，1993，3.8 节)，得到 $\{B_t\}$ 的谱密度函数由

$$S_{B_t}(f)=\sigma_X^2 C_H\sum_{j=-\infty}^{\infty}\frac{1}{|f+j|^{2H+1}},\quad -\frac{1}{2}\leqslant f\leqslant\frac{1}{2} \tag{280b}$$

给出，因此对于小的 f，$S_{B_t}(f)\propto|f|^{-1-2H}$ 近似成立，而 $\{X_t\}$ 的谱密度函数由

$$S_X(f)=4\sigma_X^2 C_H\sin^2(\pi f)\sum_{j=-\infty}^{\infty}\frac{1}{|f+j|^{2H+1}},\quad -\frac{1}{2}\leqslant f\leqslant\frac{1}{2}$$

给出(Sinai，1976；Beran，1994，p. 53，式(2.17))．对后一个结果，注意 $X_t=B_{t+1}-B_t$，所以，除了为了符合下标的习惯而进行无关紧要的单位平移外，$\{X_t\}$ 是 $\{B_t\}$ 的一阶差分．对于差分滤波器的平方增量函数是 $\mathcal{D}(f)\equiv4\sin^2(\pi f)$(参见练习[105b])，并且结果由式(268b)得到．利用欧拉-麦克劳林(Euler-Maclaurin)求和公式(Dahlquist and Björck，1974)，我们就可以利用有限的 $2M+3$ 项的求和来逼近上面的无限项求和：

$$S_X(f)\approx4\sigma_X^2 C_H\sin^2(\pi f)\Bigg(\sum_{j=-M}^{M}\frac{1}{|f+j|^{2H+1}}+\sum_{l=-1,1}\Bigg[\frac{1}{2H(lf+M+1)^{2H}}-\frac{(2H+1)(2H+2)(2H+3)}{720(lf+M+1)^{2H+4}}+\frac{2H+1}{12(lf+M+1)^{2H+2}}+\frac{1}{2(lf+M+1)^{2H+1}}\Bigg]\Bigg)$$

(详情参见 Percival 等，2000a)．在实际中，取 $M=100$ 产生足够的精度．对于较小的 f，我们近似地有 $S_X(f)\propto|f|^{1-2H}$，所以当 $-1<1-2H<0$(也就是当 $1/2<H<1$)的时候，分形高斯噪声满足式(279a)中平稳长记忆过程的定义.

图 282 最上面一行用线性/对数和对数/对数(左右两个图分别绘出)两种坐标轴展示了 $H=0.55$，0.75，0.90，0.95 的 $S_X(\cdot)$ 的图像．因为对于小的 f，$S_X(f)\propto f^{1-2H}$，当 $1/2<H<1$ 并且 H 增加的时候低频成分显著地增加，这种模式与图中的曲线是一致的．在 $S_X(f)\propto f^{1-2H}$ 近似成立的正频率上，我们有

$\log_{10}(S_X(f)) \propto (1-2H)\log_{10}(f)$，所以右边描绘出的曲线的线性状态表明对于低频，大约小于 $f=0.2$，逼近是十分好的.

图 283 中第一和第四个图分别展示了 $H=0.55$ 和 $H=0.95$ 的分形高斯噪声的实现(这些仿真的时间序列是使用 Davies-Harte 方法建立的——参见 7.8 节). 注意与它们的谱密度函数的性质相符合：与 $H=0.55$ 的时间序列相比较，$H=0.95$ 的时间序列具有成比例较强的低频成分.

当 $H=1/2$ 的时候，分形高斯噪声就简化成了白高斯噪声(即这种情况下，式(279b)具有式(268d)的形式)；当 $0<H<1/2$ 的时候，过程$\{X_t\}$的特点是具有高频振荡并且低频部分的能量不足(对这些过程，实际上有 $S_X(0)=0$).

• 完全幂规律过程

我们称离散参数的随机过程$\{X_t\}$是完全幂规律(PPL)过程，如果它的谱密度函数具有如下形式：

$$S_X(f) = C_S \mid f \mid^{\alpha}, \quad -\frac{1}{2} \leqslant f \leqslant \frac{1}{2},$$

其中 $C_S>0$. 这些过程可以分为两类，分别对应着平稳过程和非平稳过程. 第一类对应着 $-1<\alpha$ 的情况，过程是平稳的，并且它的谱密度函数具有标准意义. 对应的自协方差序列可以通过下面的公式得到：

$$s_{X,\tau} = \int_{-1/2}^{1/2} S_X(f) e^{i2\pi f\tau} df = 2C_S \int_0^{1/2} f^{\alpha} \cos(2\pi f\tau) df,$$

一般地说，这个式子没有简单的显式表达式，所以只能通过数值积分得到. 对于 $-1<\alpha<0$，完全幂规律过程显然服从式(279a)的定义，对于平稳长记忆过程——四个这类过程($\alpha=-0.1$，-0.5，-0.8 和 -0.9)的谱密度函数绘制在图 282 的中间一行. $\alpha=-0.1$ 和 $\alpha=-0.9$ 时，完全幂规律过程的仿真序列在图 283 中绘出(第 3 和第 5 幅图像). 对于 $\alpha=0$，完全幂规律过程变成了方差为 C_S 的白噪声过程. $\alpha>0$ 时的完全幂规律过程都是平稳过程，这些过程的低频成分不足，这可以从 $S_X(0)=0$ 看出来.

完全幂规律过程的第二类应用是当 $\alpha\leqslant-1$ 的时候. 这类完全幂规律过程可以理解为一类可以通过适当的差分算子变成平稳过程的非平稳过程. 在 7.7 节中给出了这方面的细节，其中解释了它们的谱密度函数的意义. 因此，完全幂规律过程在整个实数轴上都有定义.

对实数轴上一些 α，满足式(279a)的过程常常称为“$1/f$ 型过程”，所以平稳长记忆过程是 $-1<\alpha<0$ 的 $1/f$ 型过程，同时 $\alpha\leqslant-1$ 的 $1/f$ 型过程有时称作非平稳长记忆过程.

• 分形差分过程

对于平稳长记忆过程$\{X_t\}$的一个很流行的模型是分形差分(FD)过程，这个过程是由 Granger and Joyeux(1980)与 Hosking(1981)分别独立提出的. 这里$\{X_t\}$与均值零、方差 σ_ε^2(典型的高斯过程)的白噪声过程$\{\varepsilon_t\}$联系在一起：$(1-B)^{\delta}X_t=\varepsilon_t$，

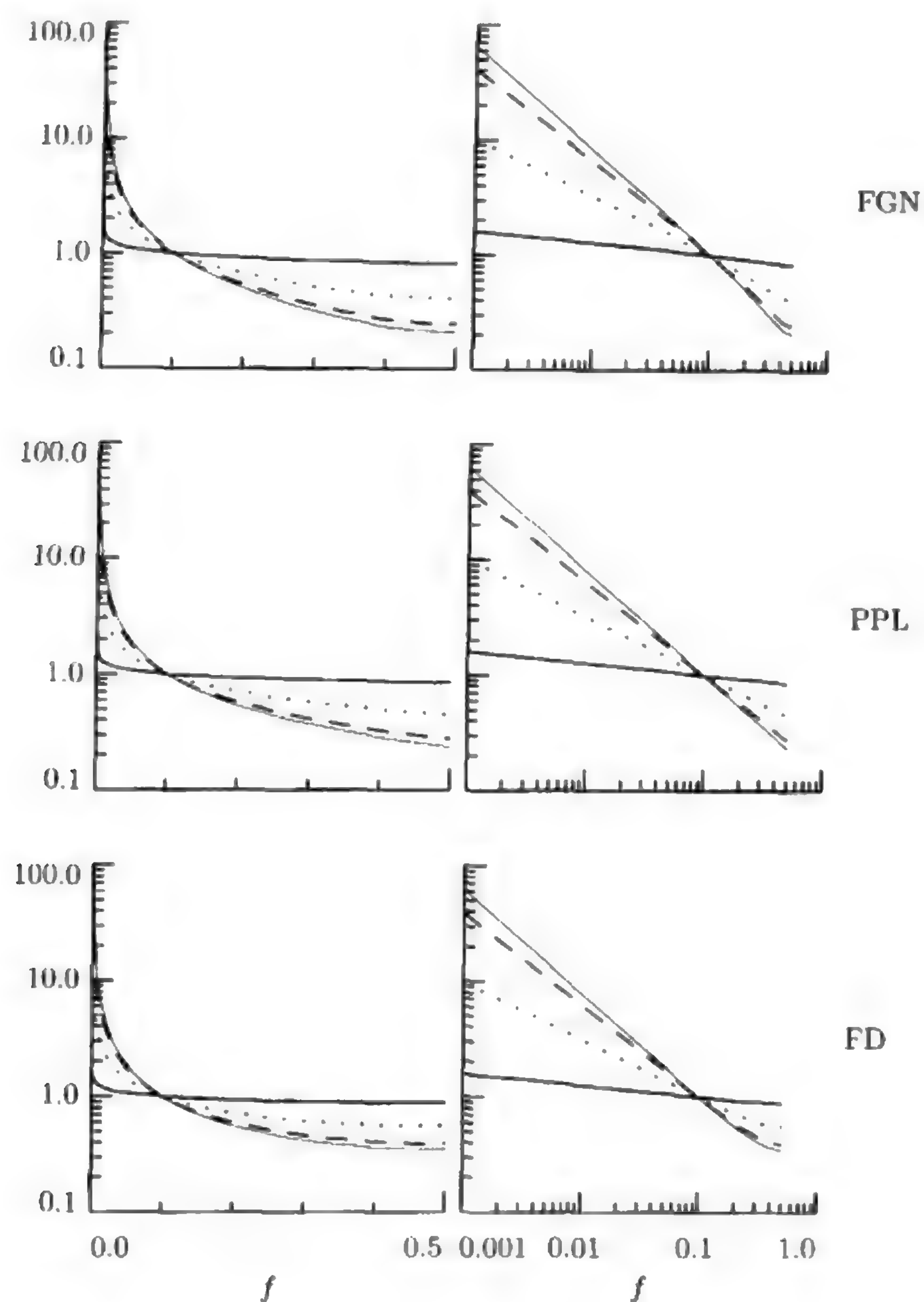


图 282 分形高斯过程、完全幂规律过程和分形差分过程的谱密度函数(分别对应着从上到下每一行)分别用线性/对数和对数/对数坐标绘制(分别对应着左右两列). 每个谱密度函数 $S_X(\cdot)$是规范化的, 满足 $S_X(0.1)=1$. 下面的表给出了所绘制的曲线的各个参数

过程	粗实线	点线	短线	细实线
FGN	$H=0.55$	$H=0.75$	$H=0.90$	$H=0.95$
PPL	$\alpha=-0.1$	$\alpha=-0.5$	$\alpha=-0.8$	$\alpha=-0.9$
FD	$\delta=0.05$	$\delta=0.25$	$\delta=0.40$	$\delta=0.45$

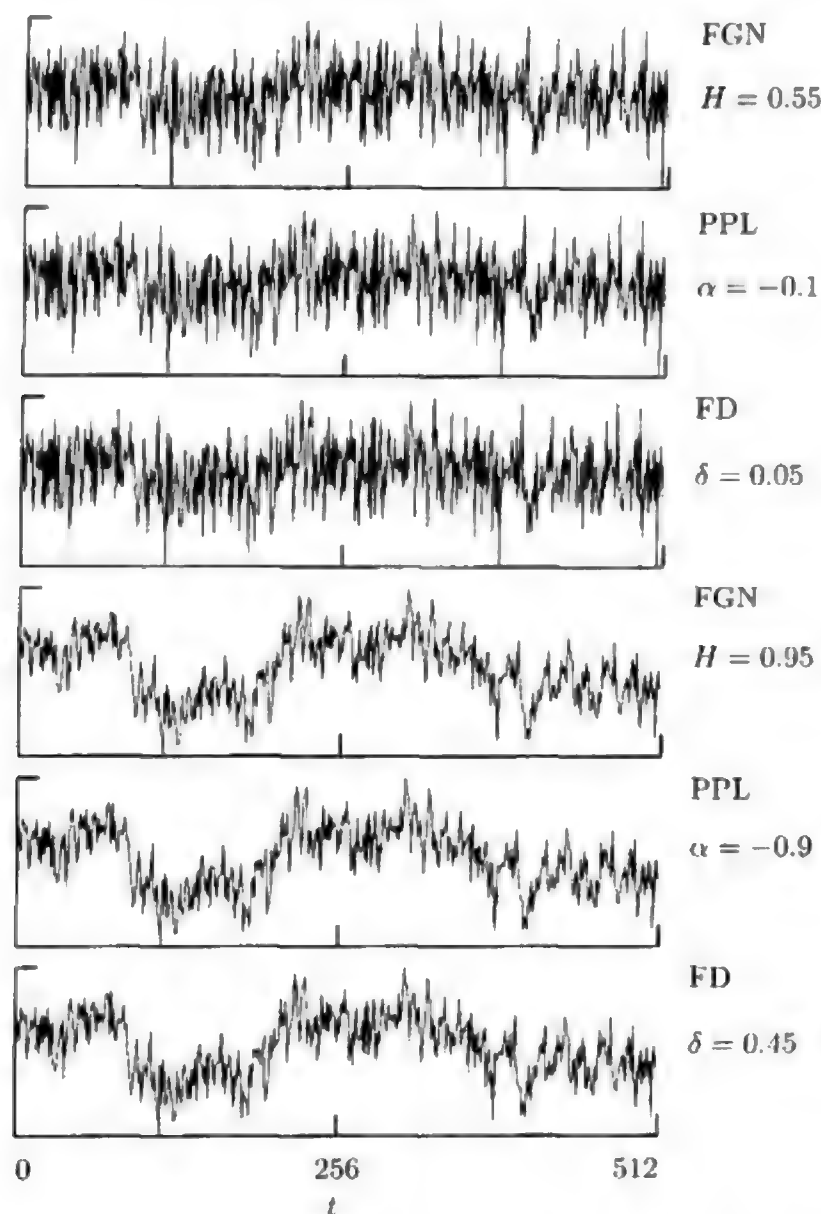


图 283 分形高斯过程、完全幂规律过程和分形差分过程仿真实现. 图 282 中的粗(细)线展示了上面(下面)三个时间序列的谱密度函数，这些谱密度函数只有在高频部分是明显的. 我们使用 Davies-Harte 方法(参见 7.8 节)形成仿真 X_0，…，X_{511}，这个方法是通过变换白噪声过程的部分，Z_0，…，Z_{1023}，实现获得的(Z_t 的值在本书的网页上给出). 为了通过可比较的 H，α，δ 说明分形高斯过程、完全幂规律过程和分形差分过程的相似性，我们使用同样的 Z_t 来创建所有这六个时间序列. 尽管上面(下面)三个时间序列看起来是相同的，它们谱密度函数估计显示出高频部分的差异，这与理论上的谱密度函数是一致的

其中 B 是向后平移算子(例如，$(1-B)X_t = X_t - X_{t-1}$)；$-1/2<\delta<1/2$；并且 $(1-B)^\delta$解释为

$$(1-B)^\delta = \sum_{k=0}^{\infty}\binom{\delta}{k}(-1)^k B^k,$$

所以

$$\sum_{k=0}^{\infty}\binom{\delta}{k}(-1)^k X_{t-k} = \epsilon_t,$$

其中

$$\binom{\delta}{k} \equiv \frac{\delta!}{k!(\delta-k)!} = \frac{\Gamma(\delta+1)}{\Gamma(k+1)\Gamma(\delta-k+1)}$$

(Beran，1994，2.5 节)．对分形差分过程的谱密度函数由

$$S_X(f) = \frac{\sigma_\varepsilon^2}{[4\sin^2(\pi f)]^\delta}, \quad -\frac{1}{2} \leqslant f \leqslant \frac{1}{2} \tag{284a}$$

给出．对于小的 f，我们近似地有 $S_X(f) \propto |f|^{-2\delta}$，所以当 $-1<-2\delta<0$(即 $0<\delta<1/2$)时，分形差分过程是平稳长记忆过程．我们使用短记号“FD(δ)”表示带有参数的分形差分过程．当 $\delta>0$ 的时候，我们有时候称 δ 为分形差分过程的“长记忆参数”．

图 282 中最下面一行展示了四个分形差分过程($\delta=0.05$，0.25，0.40，0.45)的谱密度函数．从式(284a)可以看到，谱密度函数的对数与 $|\sin(\pi f)|$ 的对数是线性相关的．事实上，因为对于小的 f，$\sin(\pi f) \approx \pi f$，在许多频率的倍频程上频率 f 的对数也是近似线性相关的．图 282 表明：当 $f>0.1$ 时，这种线性关系就破坏了，但是对于 $f \in (0, 0.1]$，逼近是非常好的．因之，这个模型与 $1/f$型过程的框架相吻合．参数 $\delta=0.05$ 和 $\delta=0.45$ 的分形差分过程的仿真序列在图 283 中的第三和最后一幅图中绘制出来．

当 $\delta=0$ 的时候，分形差分过程变成了白噪声过程；当 $-\frac{1}{2}<\delta<0$ 的时候，分形差分过程低频部分的能量是不足的，因为有 $S_X(0)=0$．当 $\delta \leqslant -\frac{1}{2}$ 的时候，我们还能用式(284a)定义平稳过程，并且当 $\delta \geqslant \frac{1}{2}$ 的时候，我们可以定义非平稳长记忆过程(对于后者，参见 7.7 节)．通过这种扩展，我们可以说分形差分过程的参数 δ 在整个实数轴上，当 $\delta<\frac{1}{2}$ 的时候保持平稳性质，当 $\delta>0$ 的时候保持长记忆性质．

因为分形差分过程是由两个参数 δ 和 σ_ε^2 完全确定的，并且因为 σ_ε^2 不决定谱密度函数的水平和它的形状，我们把具有上面谱密度函数的过程称为 FD(δ)过程．对于参数 δ 不是整数的任一平稳 FD(δ)过程的自协方差序列由

$$s_{X,\tau} = \frac{\sigma_\varepsilon^2 \sin(\pi\delta)\Gamma(1-2\delta)\Gamma(\tau+\delta)}{\pi\Gamma(\tau+1-\delta)} \tag{284b}$$

给出．置 $\tau=0$，得到了过程的方差 $s_{X,0}=\mathrm{var}\{X_t\}$，此时上式简化为

$$s_{X,0} = \frac{\sigma_\varepsilon^2\Gamma(1-2\delta)}{\Gamma^2(1-\delta)}. \tag{284c}$$

这个式子对所有的平稳$\left(\text{即当 } \delta<\frac{1}{2}\right)$分形差分过程都成立．一旦 $s_{X,0}$确定，自协方差序列的剩余的项可以简单地通过下式迭代计算：

$$s_{X,\tau} = s_{X,\tau-1}\frac{\tau+\delta-1}{\tau-\delta}, \quad \tau=1,2,\cdots \tag{284d}$$

(上式对所有的 $\delta<\frac{1}{2}$ 成立；自协方差序列向负的延伸可以根据 $s_{X,-\tau}=s_{X,\tau}$得到，这个式子对任意实值平稳过程都是成立的)．

分形高斯噪声、完全幂规律和分形差分过程都同样依赖于两个参数，其中一个控制当 $f \to 0$ 的时候逼近的幂指数；另一个参数是过程的方差(或等价地谱密度函数的水平). 这些参数的适当选择产生了分形高斯噪声、完全幂规律过程和分形差分过程，对于较小的 f 它们的谱密度函数事实上是不可区分的. 尽管有这些相似性，分形差分过程有若干优点：

[1]对于分形差分过程的谱密度函数和自协方差序列都容易计算.

[2]有一个分形差分过程的自然扩展包括具有平稳向后差分的非平稳过程(参见 7.7 节).

[3]一个分形差分过程可以认为是自回归、分形整合、移动平均(ARFIMA)过程的特例. 这个模型具有更多的参数，从而模型具有更大的自由，允许时间序列的高频部分具有更多的内容.

分形差分过程的性质和它们的扩展的充分讨论可以参见 Granger and Joyeux (1980)及 Hosking(1981)的研究文章和 Beran(1994)的书.

表 286 中总结了分形高斯噪声、完全幂规律过程和分形差分过程的参数范围，上述过程的形式为：(a)平稳长记忆过程，(b)白噪声过程，(c)平稳过程(但不是长记忆的)，(d)非平稳长记忆过程. 自相似(Hurst)系数 H，频谱斜率 α 和分形差分参数 δ 的关系在图 286 中说明，其中为了区分离散分形布朗运动中和分形高斯噪声中的两个参数 H，以后我们用 H_B 表示前者，用 H_G 表示后者. 它们之间的关系通过式(279a)得到，这个式子让我们将 H 和 δ 与 α 联系起来.

7.6 节的评论与扩展

[1]分形高斯噪声$\{X_t\}$可以通过连续参数的随机过程，著名的分形布朗运动$\{B_H(t)：0 \leqslant t < \infty\}$的抽样和差分得到. 注意 X_t 中的"t"只是无单位的指标，表示$\{X_t\}$的第 t 个元素；而 $B_H(t)$中的"t"是具有物理意义的单位(例如秒、年等等). 所以，当我们创建了抽样过程 $B_t \equiv B_H(t)$，并且做差分取形式 $X_t = B_{t+1} - B_t$ 后，$\{B_t\}$和$\{X_t\}$的抽样间隔 Δt 都等于单位长度，所以 Δt 的单位就同 $B_H(t)$中对"t"一样的单位(例如，当 $B_H(t)$中 t 的单位是秒的时候 $\Delta t = 1\text{s}$).

[2]应当注意，我们这里定义的平稳完全幂规律过程是离散参数过程$\{X_t：t = \cdots, -1, 0, 1, \cdots\}$，这就部分地意味着它们的谱密度函数定义在 $|f| \leqslant 1/2$ 上(在我们的假设下抽样步长 Δt 是单位长度，所以奈奎斯特频率是 1/2). 连续参数的完全幂规律过程$\{X(t)：-\infty < t < \infty\}$在文献上更经常提到，并且这些过程的谱密度函数定义在整个实数轴上. 这些过程应该小心地处理. 例如，注意到，对于 $-1 < \alpha < 0$，因为随着 $|f| \to \infty$，$|f|^\alpha$ 缓慢地衰减趋于零，我们有

$$\int_{-\infty}^{\infty} |f|^\alpha \mathrm{d}f = \infty.$$

表 286

过程	非平稳、长记忆	平稳、长记忆	白噪声	平稳、非长记忆
FGN	—	$\frac{1}{2}<H<1$	$H=\frac{1}{2}$	$0<H\leqslant\frac{1}{2}$
PPL	$\alpha\leqslant-1$	$-1<\alpha<0$	$\alpha=1$	$\alpha\geqslant0$
FD	$\delta\geqslant\frac{1}{2}$	$0<\delta<\frac{1}{2}$	$\delta=0$	$\delta\leqslant0$

注：每个命名了的随机过程的参数范围，这些随机过程分别是(a)非平稳长记忆，(b)平稳长记忆，(c)白噪声，(d)平稳但是非长记忆.

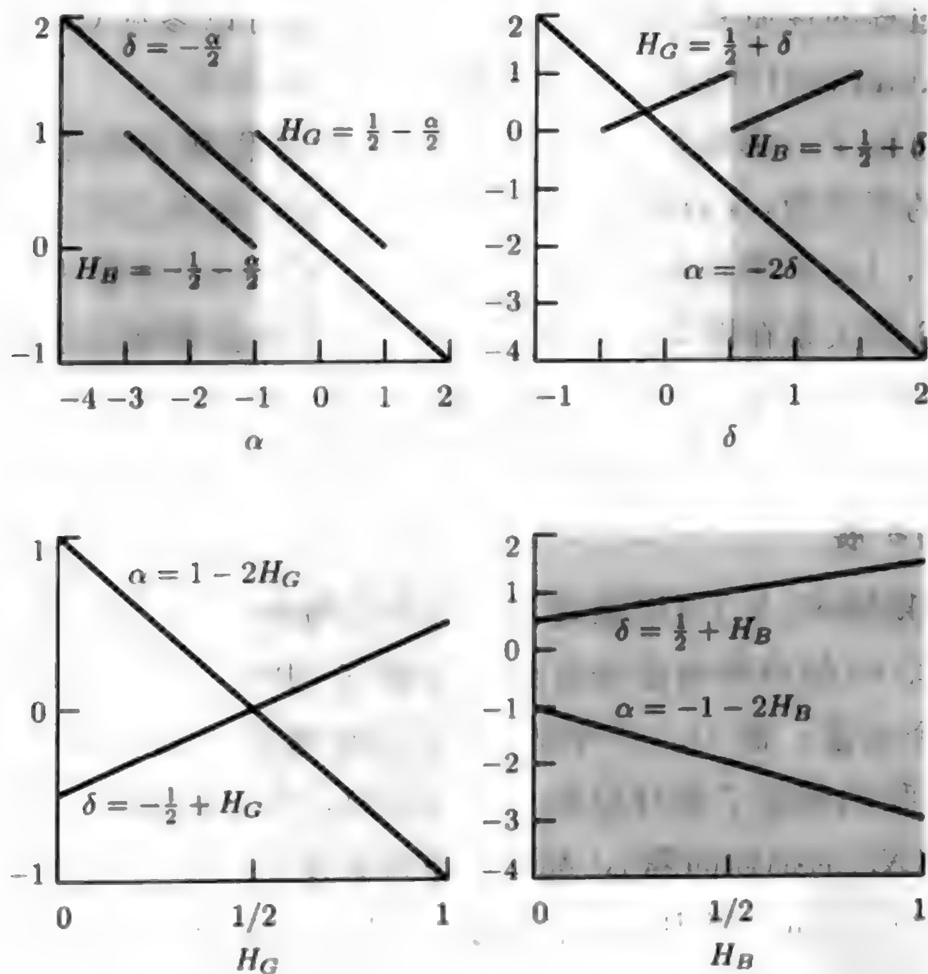


图 286 谱斜率 α、分形差分参数 δ 和 Hurst 系数 H(分形高斯过程和离散分形布朗运动)的关系. 无色的、浅色的和深色的区域表示的参数值分别对应着平稳但不是长记忆、平稳长记忆过程和非平稳长记忆过程(高斯白噪声过程发生在无色和浅色区域的交界与粗线的交叉点). 仅仅在这个图中，为了区别离散分形布朗运动和分形高斯噪声中的参数 H，在前一种情况下我们使用 H_B 但在后一种情况下使用 H_G. 记住，α 和 δ 的范围是整个实数轴，但是对于离散分形布朗运动和分形高斯噪声，都有 $0<H<1$

因为这个原因，为了获得相应于一些连续参数的平稳过程的谱密度函数，我们有

必要引入类似高频截止点的概念. 从而，假设我们定义了一个有限带宽的完全幂规律过程$\{X(t)\}$，它的谱密度函数为

$$S_{X(t)}(f)=\begin{cases}C_S\mid f\mid^{\alpha}, & \mid f\mid\leqslant f_c;\\ 0, & \text{其他},\end{cases}$$

其中$0<f_c<\infty$是高频截止点. 现在，如果定义$X_t=X(t)$，重要的是我们必须意识到除非$f_c=\frac{1}{2}$，否则不会得到一个离散参数的完全幂规律过程. 根据阶梯效应，对于$\{X_t\}$和$\{X(t)\}$的谱密度函数事实上通过下面的式子联系在一起：

$$S_X(f)=\sum_{j=-\infty}^{\infty}S_{X(t)}(f+j),\quad \mid f\mid\leqslant\frac{1}{2},$$

这一点在很多文献上被忽视了(参见下一节的评论与扩展的第[1]条).

7.7 非平稳 $1/f$ 型过程

正如 7.4 节中提到的，积分为非负有限数的任何非负偶函数$S_X(\cdot)$都可以作为某个平稳实值过程的谱密度函数. 平稳长记忆过程的谱密度函数$S_X(\cdot)$使得，当$f\to 0$时，$S_X(f)\to\infty$；而这个函数在$f=0$是奇异的，但是当$-1<\alpha<0$时，由于谱密度函数的低频部分具有$S_X(f)\propto\mid f\mid^{\alpha}$形式，所以这个函数是可积函数. 我们可以定义一类有趣的非平稳过程，它们有积分为无限值的明确定义的谱密度函数.

假设$\{X_t\}$是一个随机过程，它的n阶向后差分

$$Y_t\equiv(1-B)^d X_t=\sum_{k=0}^{d}\binom{d}{k}(-1)^k X_{t-k}$$

是谱密度函数为$S_Y(\cdot)$、均值为μ_Y的平稳过程. 这里d是非负整数，这与分形差分过程中的δ是不一样的. 与前面一样，B是向后平移算子：$BX_t=X_{t-1}$，所以$B^kX_t=X_{t-k}$. 例如我们有

$$Y_t=\begin{cases}X_t, & d=0;\\ X_t-X_{t-1}, & d=1;\\ X_t-2X_{t-1}+X_{t-2}, & d=2;\end{cases}$$

等. 如果$\{X_t\}$本身是谱密度函数为$S_X(\cdot)$的平稳过程，则由式(268b)得到的引证告诉我们$S_X(\cdot)$和$S_Y(\cdot)$通过$S_Y(f)=S_X(f)\mathcal{D}^d(f)$联系在一起，其中$\mathcal{D}(f)\equiv 4\sin^2(\pi f)$是一阶向后差分滤波器的平方增益函数；如果$\{X_t\}$是不平稳的，那么$S_X(\cdot)$通过下式定义：

$$S_X(f)\equiv\frac{S_Y(f)}{\mathcal{D}^d(f)}$$

(Yaglom，1958). 注意如果$d=0$，那么$\{X_t\}$必须是平稳的，在这种情况下$\{X_t\}$和$\{Y_t\}$是相等的. 当$\{Y_t\}$是具有参数$\sigma_\varepsilon^2>0$和$-\frac{1}{2}\leqslant\delta^{(s)}<\frac{1}{2}$的分形差分

过程时，且因此它们形如具有式(284a)右端项决定的谱密度函数，那么$\{X_t\}$的谱密度函数是

$$S_X(f)=\frac{\sigma_\varepsilon^2}{[4\sin^2(\pi f)]^{d+\delta^{(s)}}}.$$

当$d>0$时，我们可以认为上式是具有长记忆参数$\delta=d+\delta^{(s)}$非平稳分形差分过程的谱密度函数.

让我们考虑几个具有不同阶数平稳向后差分的非平稳过程的例子. 我们的第一个例子是以下述方式定义的随机游走过程. 令$\{\varepsilon_t\}$是均值为零、方差为σ_ε^2的白噪声过程，它的自协方差序列为由式(268d)给出. 注意因为$\{\varepsilon_t\}$是高斯型的和不相关的，它还是一个独立同分布序列. 利用$\{\varepsilon_t\}$，能够定义在$-\infty<t<\infty$上的随机游走过程$\{X_t\}$如下：

$$X_t\equiv\begin{cases}\sum_{u=1}^{t}\varepsilon_u, & t\geqslant 1,\\ 0, & t=0,\\ -\sum_{u=0}^{|t|-1}\varepsilon_{-u}, & t\leqslant -1.\end{cases}\tag{288a}$$

练习[288a]　证明，(a)$\{X_t\}$是非平稳过程；(b)$\{X_t\}$的一阶向后差分是均值为零的平稳过程；(c)$\{X_t\}$的谱密度函数的低频部分近似于幂规律过程，也就是说当f靠近零的时候，$S_X(f)\propto|f|^\alpha$，其中，幂规律指数α的确定也是这个练习的一部分. ◁

注意当$t\geqslant0$时，随机游走过程和参数$H=\frac{1}{2}$的离散分形布朗运动是一样的. 上式的一个变形是令$\{\varepsilon_t\}$具有某个非零均值μ_ε，此时它的一阶向后差分是均值不为零的平稳过程，而二阶向后差分是均值为零的平稳过程(参见练习7.6). 图289中第一列上面两个图分别是随机游走和修改后的随机游走的实现的例子.

为形成式(288a)的随机游走$\{X_t\}$，对零均值的过程$\{\varepsilon_t\}$累积求和. 假设对随机游走的累积求和生成了一个随机跑动过程：

$$X_t\equiv\begin{cases}\sum_{u=0}^{t}\sum_{u'=0}^{u}\varepsilon_{u'}, & t\geqslant 0,\\ 0, & t=-1,-2,\\ \sum_{u=0}^{|t|-3}\sum_{u'=0}^{u}\varepsilon_{-u'-1}, & t\leqslant -3.\end{cases}\tag{288b}$$

练习[7.7]证明了上面定义的过程是非平稳过程，这个过程的一阶和二阶向后差分分别是随机游走过程和均值为零的平稳过程. 图289中第一列中第三个图就是随机跑动过程的一个实现的例子.

作为最后一个例子，假设$\{U_t\}$是均值为零、谱密度函数为$S_U(\cdot)$的平稳过程. 构造$X_t=a_0+a_1t+U_t$，其中a_0，$a_1\neq0$为实值常数；即$\{X_t\}$是线性趋向和平稳过程的和.

练习[288b]　证明，(a)$\{X_t\}$是非平稳过程；(b)$\{X_t\}$的一阶向后差分是均值为a_1的平稳过程；(c)$\{X_t\}$的二阶向后差分是均值为零的平稳过程. ◁

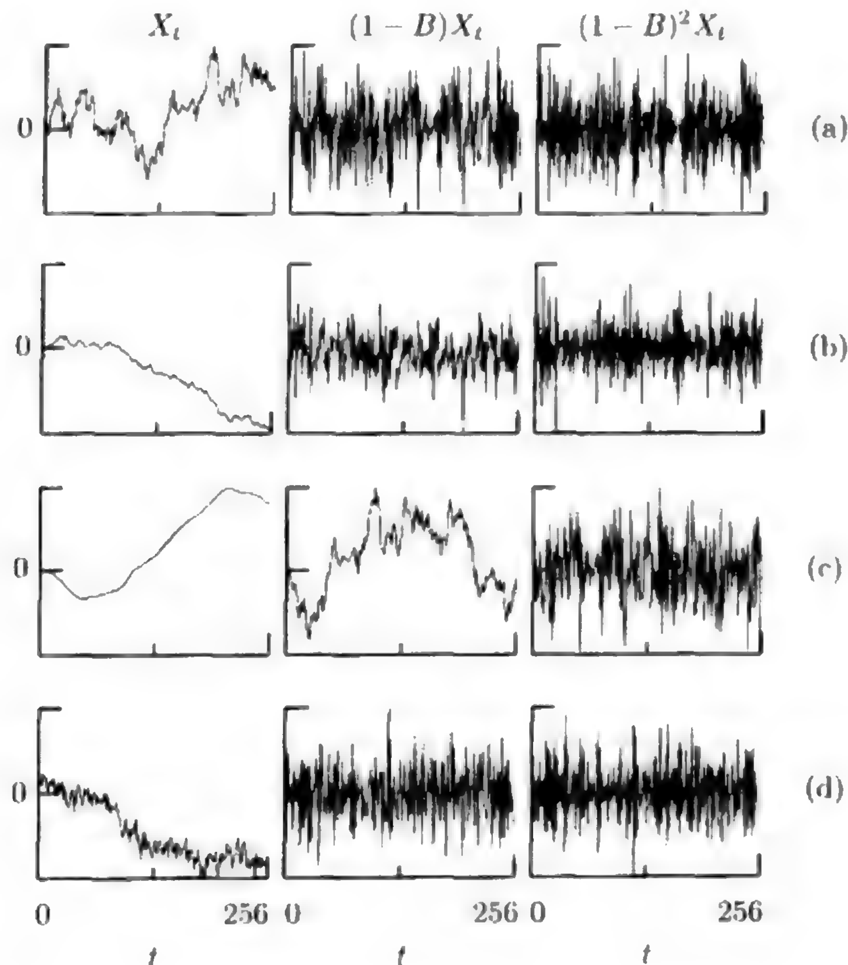


图 289 具有各阶(第一列)平稳向后差分的非平稳过程$\{X_t\}$的仿真实现与它们的一阶向后差分$\{(1-B)X_t\}$(第二列)和二阶向后差分$\{(1-B)^2X_t\}$(最后一列). 从上到下，过程分别是(a)随机游走；(b)修改的随机游走，通过具有均值$\mu_\epsilon=-0.2$的白高斯噪声序列获得；(c)随机跑动；(d)通过将线$-0.05t$和具有参数$\delta=0.45$的平稳分形差分过程的仿真求和形成的过程

利用归纳法，上面的练习可以很容易地推广，如果

$$X_t=\sum_{p=0}^{P}a_pt^p+U_t,$$

其中，$a_P\neq0$(即$\{X_t\}$是P次多项式和一个平稳过程的和)，则$\{X_t\}$的P阶向后差分是一个均值非零的平稳过程，而$P+1$阶向后差分是一个零均值的平稳过程(这是练习 7.8). 图 289 中第一列最下面一个图给出了定义为$X_t=a_0+a_1t+U_t$的过程的实现，其中$\{U_t\}$是参数$\delta=0.45$的平稳分形差分过程.

7.7 节的评论与扩展

[1]正如上一节的评论与扩展中第[2]条中所说的那样，这里我们强调的是离散参数的$1/f$型过程$\{X_t\}$而不是连续参数的过程$\{X(t)\}$. 后者的一个例子是分形布朗运动$\{B_H(t)\}$，它的谱密度函数由式(280a)给出. 注意这个谱密度函数具

有完全幂规律的形式，所以具有参数$-3<\alpha<-1$的非平稳的完全幂规律过程可以认为是$H=-(\alpha+1)/2$的分形布朗运动. 如果我们对分形布朗运动抽样获得了离散分形布朗运动，所得到的过程具有由式(280b)给出的谱密度函数；按照阶梯效应，这不是离散参数的完全幂规律过程的谱密度函数. 这个简单的事实说明离散分形布朗运动的谱密度函数对应着一个完全幂规律过程，但是这个事实却常常被很多文献忽略.

7.8 平稳随机过程的仿真

假设有一个用它的谱密度函数$S_X(\cdot)$或自协方差序列$\{s_{X,\tau}\}$描述的二阶零均值特征的高斯平稳过程$\{X_t\}$，并且我们希望创建一个时间序列，使得它是这个过程的一个实现. 如果我们能在计算机上生成一列伪随机数，使得能够认为是标准高斯随机变量的独立实现，我们就可以使用本节中描述两种方法中的一种实现这个目的. 第一种方法基于自协方差序列，并且是一种“精确”的方法，但是这种方法不是通用的方法(对某些过程，这种方法实现不了)；第二种方法基于谱密度函数，这是一种逼近方法，但是让我们感兴趣的是，对于大部分情况，逼近都是十分精确的.

我们将“精确”的方法称为Davies-Harte方法(DHM)，因为它是1987年Davies和Harte在一个统计文献上提出的(Davies and Harte，1987)；然而，事实上，这个方法在他们提出以前就已被熟知(参见1994年Johson的讨论和参考文献). Wood and Chan(1994)研究了Davies-Harte方法，并且认为这个算法在某种意义上来说是“精确”的(他们的核心论据参见练习7.10的注解). 方法假设$\{X_t\}$的自协方差序列很容易得到(例如，对分形高斯噪声和分形差分过程也是正确的；但是如果不用数值积分，这一点对完全幂规律过程是不正确的). 为了从这个过程仿真X_0，X_1，…，X_{N-1}的实现，我们必须采取下面的步骤：

[1]令$M=2N$，并且计算实值序列

$$S_k\equiv\left(\sum_{\tau=0}^{\frac{M}{2}}s_{X,\tau}\mathrm{e}^{-\mathrm{i}2\pi f_k\tau}+\sum_{\tau=\frac{M}{2}+1}^{M-1}s_{X,M-\tau}\mathrm{e}^{-\mathrm{i}2\pi f_k\tau}\right),$$

其中$f_k\equiv k/M$，$k=0$，1，…，$\frac{M}{2}$. 注意S_k能够从下面长为M的序列的离散傅里叶变换得到：

$$s_{X,0},s_{X,1},\cdots,s_{X,\frac{M}{2}-1},s_{X,\frac{M}{2}},s_{X,\frac{M}{2}-1},s_{X,\frac{M}{2}-2},\cdots,s_{X,1}$$

(参见式(36g)). 如果想要的样本尺寸是2的幂，则这个离散傅里叶变换可以通过快速傅里叶变换算法(FFT)快速计算(如果我们愿意，当$N=57$的时候，我们可以仿真一个长度为64的序列，然后丢掉前面或后面七个数).

[2]对于所有的k，验证$S_k\geqslant 0$. Gneiting(2000)证明了，参数H满足$\frac{1}{2}\leqslant H<1$的分形高斯噪声总是满足这个条件；它的证明对参数为$0\leqslant\delta<\frac{1}{2}$的分形差分过程也成立；对其他的过程，这个非负性条件也许并不满足，此时，不能用这种方法.

[3]令 Z_0，…，Z_{M-1} 是零均值、单位方差的一列（M 个）独立同分布的高斯随机变量．计算复值序列

$$\mathcal{Y}(f_k) \equiv \begin{cases} Z_0\sqrt{MS_0}, & k=0; \\ (Z_{2k-1}+\mathrm{i}Z_{2k})\sqrt{\frac{M}{2}S_k}, & 1\leqslant k<\frac{M}{2}; \\ Z_{M-1}\sqrt{MS_{\frac{M}{2}}}, & k=\frac{M}{2}; \\ \mathcal{Y}^*(f_{M-k}), & \frac{M}{2}<k\leqslant M-1. \end{cases}$$

（这里，* 表示复共轭．）

[4]最后，用离散傅里叶逆变换计算实值序列

$$Y_t \equiv \frac{1}{M}\sum_{k=0}^{M-1}\mathcal{Y}(f_k)\mathrm{e}^{\mathrm{i}2\pi f_k t}, \quad t=0,\cdots,M-1, \tag{291}$$

那么，Y_0，…，Y_{N-1} 就是 X_0，…，X_{N-1} 的仿真．

Hosking(1984)描述了另外一种仿真分形差分过程和其他高斯平稳过程的精确方法．这种方法也假设自协方差序列很容易得到，并且有对所有的高斯平稳过程都可使用的优点，但是这种算法的复杂度是 $O(N^2)$（作为比较，对于 Davies-Harte 方法中执行快速傅里叶变换的复杂度为 $O(N\log_2(N))$），因此计算速度很慢，即使是对 $N=1\ 024$ 这样大小的样本容量．

当非负性条件[2]不满足，或者谱密度函数容易但是自协方差序列无法得到的时候，我们可以使用基于离散傅里叶变换的近似仿真算法，仿真平稳高斯过程（参见 Percival，1992，详见其中参考文献）．我们称它为高斯频谱合成方法(GSSM)．假设谱密度函数 $S_X(\cdot)$ 给定，并且在整个频率域上都是有限值．为了生成序列，令 M 是正偶数．从高斯白噪声（独立同分布）序列 $\{Z_0, \cdots, Z_{M-1}\}$ 构造

$$\mathcal{Z}(f_k) \equiv \begin{cases} Z_0\sqrt{MS_X(0)}, & k=0; \\ (Z_{2k-1}+\mathrm{i}Z_{2k})\sqrt{\frac{M}{2}S_X(f_k)}, & 1\leqslant k<\frac{M}{2}; \\ Z_{M-1}\sqrt{MS_X\left(\frac{1}{2}\right)}, & k=\frac{M}{2}; \\ \mathcal{Z}^*(f_{M-k}), & \frac{M}{2}<k\leqslant M-1. \end{cases}$$

最后，利用离散傅里叶逆变换算法，计算

$$Y_t \equiv \frac{1}{M}\sum_{k=0}^{M-1}\mathcal{Z}(f_k)\mathrm{e}^{\mathrm{i}2\pi f_k t}, \quad t=0,\cdots,M-1.$$

当 M 比较大且样本容量 $N\ll M$ 时，这种逼近频率域的方法给出了十分精确的仿真(Percival，1992)．置 $M=4N$ 是一种很好的选择．X_0，…，X_{N-1} 的仿真由 Y_0，…，Y_{N-1} 给出．

为了使用高斯频谱合成方法从平稳长记忆过程仿真一个序列，我们面临一个问题，因为谱密度函数在 $f=0$ 点是无限的．我们能够对此进行调整，重新定义 $\mathcal{Z}(f_0)$ 为

$$Z_0\left[\frac{M^2}{2N-1}\left(\sum_{\tau=-(N-1)}^{N-1}s_{X,\tau}-\frac{1}{M}\sum_{k=1}^{M-1}S_X(f_k)\frac{\sin([2N-1]\pi f_k)}{\sin(\pi f_k)}\right)\right]^{1/2},$$

这个式子假设自协方差序列也是可以知道的. 如果自协方差序列不知道，则上式中自协方差序列求和可以用包含谱密度函数的积分代替

$$\sum_{\tau=-(N-1)}^{N-1} s_{X,\tau} = 2\int_0^{\frac{1}{2}} \frac{\sin([2N-1]\pi f)}{\sin(\pi f)} S_X(f)\mathrm{d}f,$$

而这个等式右端可以用数值积分计算(关于高斯频谱合成算法的细节，参见 Percival 等，2000a).

7.8 节的评论与扩展

[1] d 阶平稳向后差分过程可以很容易地通过对潜在的平稳过程的累积求和来仿真. 例如，如果$\{X_t\}$是非平稳过程，它的一阶向后差分过程 $Y_t \equiv X_t - X_{t-1}$ 是平稳的，我们可以用本节讲到的方法之一生成 Y_1，…，Y_{N-1} 仿真，由此，我们形成

$$X_t = \begin{cases} \sum_{u=1}^{t} Y_u, & t = 1,\cdots,N-1 \\ 0, & t = 0. \end{cases}$$

类似地，如果$\{X_t\}$是非平稳的，而它的二阶向后差分过程 $Y_t \equiv X_t - 2X_{t-1} + X_{t-2}$ 是平稳的，我们可以首先仿真 Y_1，…，Y_{N-1}，然后构造中间序列

$$U_t \equiv \sum_{u=1}^{t} Y_u, \quad t = 1,\cdots,N-1,$$

最后形成

$$X_t = \begin{cases} \sum_{u=1}^{t} U_u, & t = 1,\cdots,N-1, \\ 0, & t = 0. \end{cases}$$

7.9 平稳自回归过程的仿真

在上一节中，我们考虑了两种根据它的谱密度函数或自协方差序列，仿真高斯平稳过程的两种方法. 这些方法涉及到频率域中的技术，一种是精确算法，另一种是逼近算法. 对一些平稳过程的类，我们可以使用精确时间域技术有效地生成仿真序列. 这里我们给出一种利用时间域方法仿真高斯平稳 AR(p)过程(详情参见 Kay，1981).

假设$\{X_t\}$是式(268e)所描述的平稳 AR(p)过程. 令 Z_0，…，Z_{N-1} 是一列(N个)独立同分布具有零均值和单位方差的高斯随机变量. 为了生成 X_0，X_1，…，X_{N-1} 的实现，我们首先计算 $p-1$ 个序列$\{\phi_{p-1,n}: n=1, \cdots, p-1\}$，$\{\phi_{p-2,n}: n=1, \cdots, p-2\}$，…，$\{\phi_{2,n}: n=1, 2\}$和$\{\phi_{1,1}\}$，对于 $k=p$，$p-1$，…，2，第一步计算

$$\phi_{k-1,n} = \frac{\phi_{k,n} + \phi_{k,k}\phi_{k,k-n}}{1-\phi_{k,k}^2}, \quad 1 \leqslant n \leqslant k-1.$$

第二步，令 $\sigma_p^2 \equiv \sigma_\epsilon^2$，构造

$$\sigma_{k-1}^2 = \frac{\sigma_k^2}{1-\phi_{k,k}^2}, \quad k = p, p-1, \cdots, 1.$$

第三步，生成 X_0，X_1，…，X_{p-1}：

$$
\begin{aligned}
X_0 &= \sigma_0 Z_0 \\
X_1 &= \phi_{1,1} X_0 + \sigma_1 Z_1 \\
X_2 &= \phi_{2,1} X_1 + \phi_{2,2} X_0 + \sigma_2 Z_2 \\
&\vdots \\
X_{p-1} &= \phi_{p-1,1} X_{p-2} + \phi_{p-1,2} X_{p-3} + \cdots + \phi_{p-1,p-1} X_0 + \sigma_{p-1} Z_{p-1}.
\end{aligned}
$$

最后，计算剩下的 $N-p$ 个数，

$$X_t = \sum_{n=1}^{p} \phi_{p,n} X_{t-n} + \sigma_p Z_t, \quad t = p, \cdots, N-1.$$

上面简要列举的 X_0，…，X_{p-1} 的生成提供了平稳过程的初始值，其余数据的生成是通过式(268e)中的 AR(p)过程的定义实现的.

7.10 练习

[7.1] 假设 X 是自由度为 η 的 χ^2 随机变量；即它具有式(263b)中的概率分布函数. 利用式(258b)确定 $Y \equiv \log(X)$ 的概率分布函数 $f_Y(\cdot\,;\,\eta)$. 绘出 $\eta=10$，12，16 时 $f_Y(y;\,\eta)$ 关于 $y(0<y\leqslant 4)$ 的函数图像，验证图 276 中细线绘出的图是正确的.

[7.2] 这里我们考虑协方差的一些基本性质：

(a)如果 X 是随机变量且 c 是常数，则 $\text{cov}\{X,\ c\}=0$.

(b)证明，$|\text{cov}\{X,\ Y\}|^2 \leqslant \text{var}\{X\}\text{var}\{Y\}$.

(c)假设 X 和 Y 都是随机变量，至少有一个的均值为零，证明：$\text{cov}\{X,\ Y\}=E\{XY\}$.

(d)给定两个随机变量的有限集合 $\{X_j\}$ 和 $\{Y_k\}$ 及两个相应的常数的集合 $\{a_j\}$ 和 $\{b_k\}$，证明

$$\text{cov}\left\{ \sum_j a_j X_j, \sum_k b_k Y_k \right\} = \sum_j \sum_k a_j b_k \text{cov}\{X_j, Y_k\}.$$

[7.3] 令 Z_1，Z_2，…，Z_η 是独立同分布零均值和单位方差的高斯随机变量. 证明：$\chi_\eta^2 \equiv Z_1^2 + Z_2^2 + \cdots + Z_\eta^2$ 具有由式(263b)给出的概率分布函数.

[7.4] 令 X_0，…，X_{N-1} 是具有自协方差序列为 $\{s_{X,\tau}\}$ 的平稳过程的一部分，令 $\overline{X} \equiv \frac{1}{N}\sum_{t=0}^{N-1} X_t$ 是样本均值. 证明

$$\text{var}\{\overline{X}\} = \frac{1}{N} \sum_{\tau=-(N-1)}^{N-1} \left(1 - \frac{|\tau|}{N}\right) s_{X,\tau}. \tag{293}$$

[7.5] 基于 7.5 节给出的周期图的分布结果，验证式(270c).

[7.6] 令 $\{\epsilon_t\}$ 是高斯白噪声过程，使得 $E\{\epsilon_t\}=\mu_\epsilon \neq 0$. 假设我们利用这个非零过程构造式(288a)描述的随机游走过程 $\{X_t\}$. 证明：

(a) $\{X_t\}$的一阶向后差分是具有非零均值的平稳过程；

(b) $\{X_t\}$的二阶向后差分是具有零均值的平稳过程.

[7.7] 令$\{X_t\}$是式(288b)定义的随机跑动过程. 证明：

(a) $\{X_t\}$是非平稳过程；

(b) $\{X_t\}$的二阶向后差分是均值为零的平稳过程；

(c) $\{X_t\}$的一阶向后差分是随机游走过程(式(288a))；

(d) $\{X_t\}$的谱密度函数的低频部分近似是一个幂规律过程，其中幂规律指数α的确定也是这个练习的一部分.

[7.8] 令$\{U_t\}$是零均值和谱密度函数为$S_U(\cdot)$的平稳过程. 证明：如果

$$X_t = \sum_{p=0}^{P} a_p t^p + U_t,$$

其中$a_P \neq 0$，那么$\{X_t\}$的P阶向后差分是均值非零的平稳过程，而$\{X_t\}$的$(P+1)$阶向后差分是均值为零的平稳过程.

[7.9] 利用随机数生成器仿真不相关的，具有零均值、单位方差的高斯(正态)分布的128个数据组成的Z_t，组成一个时间序列$\{Z_t: t=0, \cdots, 127\}$(这个时间序列可以看作是长度为128的高斯白噪声过程的实现). 计算$\{Z_t\}$的单位尺度的$D(4)$极大重叠离散小波变换小波系数$\widetilde{\mathbf{W}}_1$，然后计算并且绘出$\widetilde{\mathbf{W}}_1$的旋转累积方差序列(参见式(189)). 生成很多个的类似的白高斯序列；对于每个序列计算它们的旋转累积方差序列，并且绘出这些旋转累积方差序列的平均值关于$t=0, \cdots, 127$的图像. 如果你重复这个过程至无限多，那么绘出的旋转累积方差序列平均值的图像会收敛到什么程度呢?

[7.10] 证明，式(291)的过程$\{Y_t\}$是零均值高斯平稳过程，它的自协方差下标$\tau=0, \cdots, N-1$的项用$s_{X,0}, s_{X,1}, \cdots, s_{X,N-1}$给出(我们能够由这个结果得知随机变量$Y_0, \cdots, Y_{N-1}$与随机变量$X_0, \cdots, X_{N-1}$具有同样的统计性质).

[7.11] 根据7.9节中定义的变量，AR(p)过程的方差是多少?

第 8 章

小 波 方 差

8.0 引言

就像第 4 章和第 5 章所讲的，离散小波变换和它的方差、极大重叠离散小波变换的一个重要用途就是在尺度对尺度的基础上分解时间序列的样本方差. 在本章，通过定义一个理论上非常著名的小波方差(有时称为小波谱)，更深一步探究基于小波的方差分析(ANOVA). 基于离散小波变换或极大重叠离散小波变换，这个理论上的方差很容易估计，且已经成功地用于很多应用上；例如，参见 Gamage(1990)，Bradshaw and Spies(1992)，Flandrin(1992)，Gao and Li(1993)，Hudgins 等(1993)，Kumar and Foufoula-Georgiou(1993，1997)，Tewfik 等(1993)，Wornell(1993)，Scargle(1997)，Torrence and Compo(1998)和 Carmona 等(1998). 小波方差的定义和考虑其合理性将在 8.1 节给出，随后在 8.2 节讨论它的基本特征，在 8.3 节考虑通过给定一个具有平稳向后差分的随机过程的长为 N 的一个部分时间序列，如何估计小波方差. 我们研究小波方差估计量的大样本统计特征，并讨论基于估计小波方差上为真正的小波方差确定一个近似置信区间的方法(8.4 节). 之后，在 8.5 节描述小波方差的估计如何转化为谱密度函数(SDF)的估计. 在总结结论(8.10 节)之前，我们考虑原子钟的频率平稳性、子潮汐海平面涨落、尼罗河最低水位和海洋垂直切变，小波方差会告诉我们这四个时间序列的分析(8.6 节到 8.9 节).

8.1 小波方差的定义和理论基础

如 5.4 节，令$\{\tilde{h}_{j,l}: l=0, \cdots, L_j-1\}$是一个具有尺度 $\tau_j\equiv 2^{j-1}$ 的第 j 层极大重叠离散小波变换的小波滤波器，其中 $L_j\equiv(2^j-1)(L-1)+1$ 是滤波器的宽度(这里 $L=L_1$ 是 5.2 节描述的单位尺度滤波器$\tilde{h}_l=\tilde{h}_{1,l}$的宽度). 令$\{X_t: t=\cdots,$

$-1, 0, 1, \cdots\}$表示一个离散参数实值随机过程，并且令

$$\overline{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l}, \quad t = \cdots, -1, 0, 1, \cdots \tag{296a}$$

表示用极大重叠离散小波变换的小波滤波器$\{\tilde{h}_{j,l}\}$滤波$\{X_t\}$得到的随机过程(密切注意上面的与式(196a)定义的$\widetilde{W}_{j,t}$比较：上面包含滤波一个无穷序列，然而$\widetilde{W}_{j,t}$来自循环滤波一个有限序列). 如果它存在并且是有限的，尺度τ_j上的与时间有关的小波方差定义为$\overline{W}_{j,t}$的方差；即

$$\nu_{X,t}^2(\tau_j) \equiv \text{var}\{\overline{W}_{j,t}\}.$$

在本章，我们限制在过程$\{X_t\}$以使$\nu_{X,t}^2(\tau_j)$存在，有限且与t无关(在下一节中将给出足够的条件以使这些叙述成立). 这些限制基本上说明尺度τ_j上的$\{X_t\}$统计特征是随时间不变的，因此，可以通过与时间无关的小波方差有效地概括

$$\nu_X^2(\tau_j) \equiv \text{var}\{\overline{W}_{j,t}\}. \tag{296b}$$

虽然在发展此理论过程中，通过限制以产生与时间无关小波方差，我们将会通过8.7到8.9节的例子详细说明小波方差适用于处理有着与时间有关的小波方差的一些过程.

我们通常提供三个理由来说明为什么小波方差是有意义的. 第一，小波方差在尺度对尺度的基础上分解一些随机过程的方差，因此，相当吸引那些思考按照变化控制一系列不同尺度的物理现象的科学家. 第二，小波方差与谱密度函数的概念紧密联系，并且在这个函数有一个相当简单的结构下提供了谱密度函数的一个简单总结(在物理学科中常常是正确的). 第三，小波方差是一个过程方差的有效替代品，此过程有无穷方差；类似地，小波方差的估计量是过程样本方差的有效替代，此过程样本方差不能很好地表现样本特征. 现在让我们详细考虑每一个理由.

• *方差分解*

假设$\{Y_t\}$是一个有定义在频率区间$\left[-\frac{1}{2}, \frac{1}{2}\right]$上的谱密度函数$S_Y(\cdot)$的平稳过程. 如式(267d)所注意的(为方便起见用Δt代表单位)，谱密度函数的一个基本性质是

$$\int_{-1/2}^{1/2} S_Y(f)\,\mathrm{d}f = \sigma_Y^2 \equiv \text{var}\{Y_t\}; \tag{296c}$$

即谱密度函数通过频率分解过程方差. 对于小波方差，这个基本结果类似于

$$\sum_{j=1}^{\infty} \nu_Y^2(\tau_j) = \sigma_Y^2 \tag{296d}$$

(参见本节评论与扩展第[1]条). 因此，就像谱密度函数通过频率分解 σ_Y^2，小波方差关于一个离散无关变量 τ_j，也就是知道的尺度，分解 σ_Y^2(为了简化我们的讨论，本章大部分内容处理标准尺度 τ_j，但是当处理实际应用时，我们必须将它们转向物理尺度 $\tau_j \Delta t$). 粗略地说，$\nu_Y^2(\tau_j)$相当于过程$\{Y_t\}$的两次加权平均，第一次包含宽度为 τ_j 的区间内的值，而第二次包含围绕这个区间附近的数据(或者是在哈尔小波情况下接近它)——这些加权平均的差异越大，$\nu_Y^2(\tau_j)$就越大. 作为 τ_j 的函数的$\nu_Y^2(\tau_j)$图显示了哪些尺度对过程方差是重要的(注意以上类似离散小波经验功率谱 $P_W(\tau_j)$，它通过尺度分解样本方差——参见式(62)).

让我们专门研究小波方差的最简单的例子，即基于极大重叠离散小波变换的哈尔小波滤波器$\left\{\frac{1}{2}, -\frac{1}{2}\right\}$上的例子. 如 8.6 节所讨论的，小波方差与能很好测试原子钟性能的 Allan 方差成比例. 作为 τ_j 函数的 Allan 方差的平方根绘的图自从 1960s 就常常用于描绘钟在不同时间段是如何很好地保持时间；然而，就像将在 8.9 节中看到的，Allan 方差会误导解释某些物理过程，对于它，基于 $L \geqslant 4$ 的小波滤波器更合适(参见 Percival and Guttorp，1994).

• **谱密度函数的有效替代**

小波方差还是很有意义的，因为它提供了一种调整谱密度函数的方法. 就像在第 4 章和第 5 章中看到的，尺度 τ_j 上的小波系数与区间$\left[\frac{1}{2^{j+1}}, \frac{1}{2^j}\right]$上的频率联系在一起. 由于 $\nu_Y^2(\tau_j)$仅仅是尺度 τ_j 上的极大重叠离散小波变换小波系数的方差，在一些合理的条件下，频率和尺度的记号之间紧密联系意味着

$$\nu_Y^2(\tau_j) \approx 2\int_{1/2^{j+1}}^{1/2^j} S_Y(f)\mathrm{d}f \tag{297a}$$

(其准确联系参见式(305)——注意积分前面的因子 2 是由于谱密度函数 $S_Y(\cdot)$是区间$\left[-\frac{1}{2}, \frac{1}{2}\right]$上关于 f 的偶函数). 小波方差通过每个倍频程频带内仅仅一个值总结谱密度函数中的信息，并且当谱密度函数在每个倍频程带内相对无特点时就特别有用. 例如，假设$\{Y_t\}$是一个纯粹的幂规律过程；即它的谱密度函数由 $S_Y(f) \propto |f|^\alpha$ 给出(参见 7.6 节). 如果使用(297a)中的逼近，可以看到近似地得到

$$\nu_Y^2(\tau_j) \propto \tau_j^{-\alpha-1}. \tag{297b}$$

以 $\log(\tau_j)$为横坐标的 $\log(\nu_Y{}^2(\tau_j))$图的线性变化指出一个幂规律过程的存在性，并且直线的斜率能够用于推断幂规律的指数 α(参见 9.5 节). 对于这个简单的模型，由小波方差给出的总结没有遗漏任何信息.

有趣的是，如果我们再一次研究哈尔小波方差，Blackman and Tukey(1958，18 节)的试验谱和使用与(297a)的这个小波方差相同. 试验谱通过 Blackman 和

Tukey 推荐为一个谱密度函数的"简单和便宜"的初步估计(参见 8.5 节评论与扩展中对试验谱的更深一步讨论). 小波方差基于不同于哈尔的小波是一个有用的推广，因为对于一些过程，(297a)中的近似随着小波滤波器宽度的增加而改善.

因为小波方差是谱密度函数的一个调整，小波方差的估计比谱密度函数的非参数估计更直接. 假定目前有一个长为 N 的时间序列，它能够认为是有未知均值的平稳过程$\{Y_t\}$的一部分 Y_0，…，Y_{N-1} 的实现，如 7.5 节所描述的，离散傅里叶变换产生 $S_Y(\cdot)$的一个基本估计称为周期图，由

$$\hat{S}_Y^{(p)}(f_k) \equiv \frac{1}{N}\left|\sum_{t=0}^{N-1}(Y_t - \bar{Y})e^{-i2\pi f_k t}\right|^2$$

给出，其中 $f_k \equiv k/N$，且 $\bar{Y} \equiv \sum_{t=0}^{N-1} Y_t/N$ 是 Y_t 的样本均值. 周期图满足式(296c)的抽样方案，即

$$\frac{1}{N}\sum_{k=-\lfloor N/2\rfloor+1}^{\lfloor N/2\rfloor}\hat{S}_Y^{(p)}(f_k) = \frac{1}{N}\sum_{t=0}^{N-1}(Y_t - \bar{Y})^2$$

(这里$\lfloor x\rfloor$指的是小于等于 x 的最大整数)；然而，如 7.5 节所讨论的，周期图在有用性上有限，因为它是 $S_Y(\cdot)$的一个不一致估计，且因为它对于通常统计学家认为的大样本会严重偏离(像图 273 演示的 $N=2\ 048$). 我们必须克服周期图的简易性去获得一个有用的谱密度函数估计.

作为对比，第 5 章中描述的 Y_0，…，Y_{N-1} 的极大重叠离散小波变换直接产生在尺度 τ_j 上的小波方差的一个有用估计量，即

$$\tilde{\nu}_Y^2(\tau_j) \equiv \frac{1}{N}\sum_{t=0}^{N-1}\widetilde{W}_{j,t}^2.$$

如能够从式(171a)和练习[171b]立即看到的，上面导致式(296d)的下述抽样方案：

$$\sum_{j=1}^{J_0}\tilde{\nu}_Y^2(\tau_j) + \frac{1}{N}\sum_{t=0}^{N-1}(\widetilde{V}_{J_0,t} - \bar{Y})^2 = \frac{1}{N}\sum_{t=0}^{N-1}(Y_t - \bar{Y})^2.$$

另外，尽管周期图会严重偏离，可以在不受边界条件影响的基于 $\widetilde{W}_{j,t}$项上容易构造 $\nu_Y^2(\tau_j)$的一个无偏估计量(参见 8.3 节)；对此陈述的防止误解的一个说明，见下面评论与扩展的第[2]条. 进而，鉴于周期图是不一致的，在 8.3 节我们将为这个无偏估计量建立一致性. 对于相应的无特征谱，小波方差是一个吸引人的代替描述，易于解释和估计.

• 过程方差的有效代替

平稳过程$\{Y_t\}$的一个基本特征是它的方差σ_Y^2. 如果对于$\{Y_t\}$过程均值μ_Y已知，那么

$$\tilde{\sigma}_Y^2 \equiv \frac{1}{N}\sum_{t=0}^{N-1}(Y_t-\mu_Y)^2$$

是σ_Y^2的一个无偏估计量，因为

$$E\{\tilde{\sigma}_Y^2\} = \frac{1}{N}\sum_{t=0}^{N-1}E\{(Y_t-\mu_Y)^2\} = \sigma_Y^2.$$

另一方面，如果μ_Y是未知的，并且由样本均值$\bar{Y}$估计，方差σ_Y^2的通常估计量是

$$\hat{\sigma}_Y^2 \equiv \frac{1}{N}\sum_{t=0}^{N-1}(Y_t-\bar{Y})^2. \tag{299a}$$

练习[299] 证明

$$E\{\hat{\sigma}_Y^2\} = \sigma_Y^2 - \operatorname{var}\{\bar{Y}\}. \tag{299b}$$

◁

由于$\operatorname{var}\{\bar{Y}\}\geqslant 0$，$E\{\hat{\sigma}_Y^2\}$必定是非负的，我们一定有

$$0 \leqslant E\{\hat{\sigma}_Y^2\} \leqslant \sigma_Y^2 \tag{299c}$$

(这是极好结果的一种特殊情况，归于 David，1985). 因此，一般来说，不能用$\hat{\sigma}_Y^2$过高估计σ_Y^2. 如果随着$\tau\to\infty$，$s_{Y,\tau}\to 0$(对于具有一个谱密度函数的平稳过程通常是正确的)，时间序列分析的一个标准结果(Fuller，1996，推论 6.1.1.1)表示，随着$N\to\infty$，$\operatorname{var}\{\bar{Y}\}\to 0$，由此可以得出，随着$N\to\infty$，$E\{\hat{\sigma}_Y^2\}\to\sigma_Y^2$，即$\hat{\sigma}_Y^2$是$\sigma_Y^2$的一个渐近无偏估计.

样本方差的偏差只不过由于缺少对过程均值的了解，而$\hat{\sigma}_Y^2$的渐近无偏性将提供方便，如果样本尺寸N足够大，一切会变得很好，事实上，存在产生样本方差$\hat{\sigma}_Y^2$的平稳过程，此样本方差对于非常大样本能够是σ_Y^2的一个严重偏离估计量. 这些过程不是病态的，而对于许多通常在物理科学收集到的时间序列是非常合理的模型：事实上，它们是 7.6 节中的平稳长记忆过程. 我们可以用这些过程去支持以下的结论(看起来是不可容忍的)：对于每一个样本尺寸$N\geqslant 1$和每一个$\varepsilon>0$，存在一个平稳过程，使得

$$\frac{E\{\hat{\sigma}_Y^2\}}{\sigma_Y^2} < \varepsilon. \tag{299d}$$

总之，不管我们取的样本尺寸N是多少，总有一个平稳过程，对于它，样本方差的期望值可以是真正过程方差的任意小的比例. 为了建立这个结论，假定$\{Y_t\}$是分形高斯噪声(FGN)，由定义，它是一个有由式(279b)决定的自协方差序列的一个平稳过程，即

$$s_{Y,\tau} = \frac{\sigma_Y^2}{2}(|\tau+1|^{2H} - 2|\tau|^{2H} + |\tau-1|^{2H}), \quad \tau = \cdots,-1,0,1,\cdots,$$

其中H是 Hurst 满足$0<H<1$的参数.

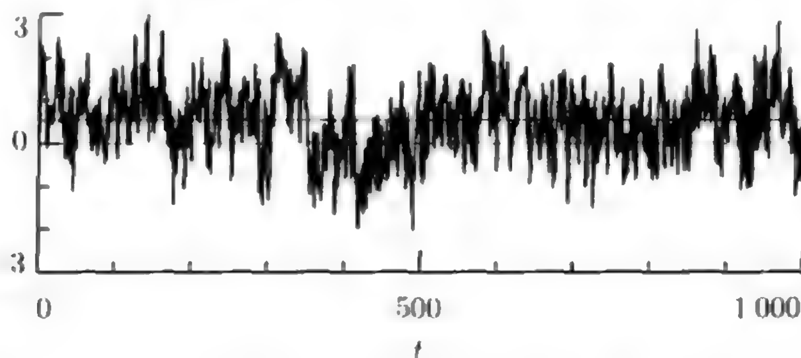


图 300 具有自相似参数 $H=0.9$ 的分形高斯噪声过程的实现. 近似样本均值为 0.53，真实的均值 0 由细的水平线表示

练习[300] 证明

$$\operatorname{var}\{\bar{Y}\} = \sigma_Y^2 N^{2H-2}. \tag{300}$$

◁

提示：使用归纳法.

由式(299b)我们有

$$\frac{E\{\hat{\sigma}_Y^2\}}{\sigma_Y^2} = 1 - N^{2H-2}.$$

对于 $\varepsilon<1$ 和 $N>1$，这个结论通过取 H，使

$$H > 1 + \frac{\log(1-\varepsilon)}{2\log(N)}$$

(情形 $\varepsilon\geqslant 1$ 直接由式(299c)得到，而 $N=1$ 是平凡的，因为此时$\hat{\sigma}_Y^2=0$).

注意对于小的 ε 和大的 N，任何使式(299d)成立的分形高斯噪声有一个 Hurst 参数仅仅小于 1. 对于一些小的 $\delta>0$，记 $H=1-\frac{\delta}{2}$，我们看到式(300)变成 $\operatorname{var}\{\bar{Y}\}=\sigma_Y^2/N^\delta$. 当$\{Y_t\}$是一个白噪声过程时，这能够与 $\operatorname{var}\{\bar{Y}\}=\sigma_Y^2/N$ 比较(对分形高斯噪声，当 $H=\frac{1}{2}$时这是真的). 因此，对于有一个 H 接近于 1 的分形高斯噪声，$\operatorname{var}\{\bar{Y}\}$减小至 0 的速度比对于无关联的观测值的$\frac{1}{N}$慢得多.

作为例子，图 300 展示了一个来自 $\mu_Y=0$，$\sigma_Y^2=1$ 和 $H=0.9$ 的分形高斯噪声过程的长为 $N=1\ 000$ 的时间序列的实现. 这个时间序列的样本均值是 $\bar{Y}\doteq 0.53$ (在图中由一条细的水平线表示). 如果我们用知道的 $\mu_Y=0$ 去估计过程方差，得到$\tilde{\sigma}_Y^2\doteq 0.99$，非常接近真实方差；另一方面，如果我们不假设知道 μ_Y 而代替使用 $\bar{Y}$，我们得到$\hat{\sigma}_Y^2\doteq 0.71$，它低于估计的真实值 σ_Y^2，且接近于我们期望得到的，因为 $E\{\hat{\sigma}_Y^2\}\doteq 0.75$. 事实上，对于这个特殊的分形高斯噪声过程，我们需要 $N=10^{10}$或更大的样本，以使 $\sigma_Y^2-E\{\hat{\sigma}_Y^2\}\leqslant 0.01$，即为了使真实方差的偏差为 1 或小于 1.

从此讨论获得的信息，由于用样本均值估计过程均值的必要性，对于一些平

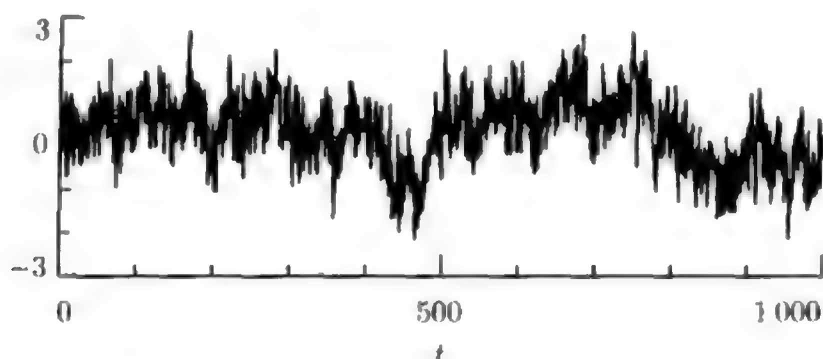


图 301 一个随机过程的实现，它的一次向后差分是一个有着自相似参数$H=0.1$的分形高斯噪声过程. 根据它的确切定义，这个过程的方差要么对于所有的t是无穷，要么依赖于时间，随着$|t|\to\infty$增加至无穷

稳过程，样本方差具有很小的偏离特性(当一个平稳过程用于作为一个可观测到的时间序列的模型时，μ_Y难得先验知道). 另外，也有一些具有平稳差分的不平稳过程(参见 7.7 节)，对于它们，样本方差不是一个特别有用的统计量，因为一个过程方差不能定义得既有限又是时间不变的. 尽管如此，由这些过程实现有限部分，能够像一些在物理科学中可以观测到的时间序列且粗略看起来像平稳过程的实现(参见图 301)!

小波方差对平稳过程来说(和有着平稳向后差分的过程)通常是一个有用的概念，对于这些过程样本方差的效用值得怀疑. 基本观点是用在某一尺度上通过样本方差为可变性作全局观测来代替可变性的概念. 为了了解为什么小波方差是一个比过程方差更容易估计的量，再一次假设$\{Y_t\}$是一个有着未知均值μ_Y的平稳过程. 由于随机过程$\{\overline{W}_{j,t}\}$是通过用极大重叠离散小波变换小波滤波器$\{\tilde{h}_{j,l}\}$滤波$\{Y_t\}$得到的，由此得到$\{\overline{W}_{j,t}\}$是一个过程方差是小波方差$\nu_Y^2(\tau_j)$的平稳过程(参见式(268a)和(267d)附近的讨论)，然而$\{\overline{W}_{j,t}\}$的过程均值是已知的，因为

$$E\{\overline{W}_{j,t}\} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} E\{Y_{t-l}\} = \mu_Y \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} = 0$$

(参见表 202). 因为$\{\overline{W}_{j,t}\}$的过程均值已知为 0(本质上是由于嵌入在小波滤波器中的差分格式)，容易得到它的过程方差的无偏估计量，即小波方差$\nu_Y^2(\tau_j)$的一个无偏估计量.

8.1 节的评论与扩展

[1]这里，我们证明对于具有谱密度函数$S_Y(\cdot)$的任一平稳过程$\{Y_t\}$，式(296d)成立. 首先，我们定义

$$\overline{V}_{J_0,t} \equiv \sum_{l=0}^{L_{J_0}-1} \tilde{g}_{J_0,l} Y_{t-l},$$

它是使用极大重叠离散小波变换尺度滤波器$\{\tilde{g}_{J_0,l}\}$滤波$\{Y_t\}$得到的平稳过程. 现在我们要求，对于任一整数$J_0\geqslant 1$，

$$\mathrm{var}\{Y_t\}=\mathrm{var}\{\overline{V}_{J_0,t}\}+\sum_{j=1}^{J_0}\nu_Y^2(\tau_j). \tag{302}$$

因为$\{\overline{W}_{j,t}\}$是通过用第j层极大重叠离散小波变换小波滤波器$\{\tilde{h}_{j,l}\}$滤波平稳过程$\{Y_t\}$得到的，式(268b)告诉我们，$\{\overline{W}_{j,t}\}$是有着由$S_j(f)\equiv\widetilde{\mathcal{H}}_j(f)S_Y(f)$定义的谱密度函数的一个平稳过程，其中$\widetilde{\mathcal{H}}_j(\cdot)$是$\{\tilde{h}_{j,l}\}$的平方增益函数. 由于小波方差$\nu_Y^2(\tau_j)$是$\{\overline{W}_{j,t}\}$的方差且谱密度函数$S_j(\cdot)$的积分等于这个方差，我们有

$$\nu_Y^2(\tau_j)=\int_{-1/2}^{1/2}\widetilde{\mathcal{H}}_j(f)S_Y(f)\,\mathrm{d}f;$$

类似地，由于$\widetilde{\mathcal{G}}_{J_0}(\cdot)$是$\{\tilde{g}_{J_0,l}\}$的平方增益函数，

$$\mathrm{var}\{\overline{V}_{J_0,t}\}=\int_{-1/2}^{1/2}\widetilde{\mathcal{G}}_{J_0}(f)S_Y(f)\,\mathrm{d}f.$$

使用式(163c)，对于所有的f，我们有$\widetilde{\mathcal{G}}_1(f)+\widetilde{\mathcal{H}}_1(f)\equiv\widetilde{\mathcal{G}}(f)+\widetilde{\mathcal{H}}(f)=1$，由此我们得到

$$\mathrm{var}\{Y_t\}=\int_{-1/2}^{1/2}S_Y(f)\,\mathrm{d}f=\int_{-1/2}^{1/2}[\widetilde{\mathcal{G}}_1(f)+\widetilde{\mathcal{H}}_1(f)]S_Y(f)\,\mathrm{d}f=\mathrm{var}\{\overline{V}_{1,t}\}+\nu_Y^2(\tau_1),$$

因此对于$J_0=1$式(302)成立.

练习[302] 通过证明，如果$J_0\geqslant 2$且如果对于J_0-1式(302)成立，就必定还对J_0成立，完成结论的证明. 提示：研究表202. ◁

如果我们可以证明随着$J_0\to\infty$，$\mathrm{var}\{\overline{V}_{J_0,t}\}\to 0$，式(296d)的证明从(302)得到. 直观地，我们希望这是真的，因为$\{\overline{V}_{J_0,t}\}$是一个有由$|f|\in[0, 1/2^{J_0+1}]$上给出标称通带的$\{Y_t\}$的一个低通滤波的结果，它的带宽随着$J_0\to\infty$减至0. 为了组织形式上的一个证明，我们必须证明，对于任意的$\varepsilon>0$，存在一个J_ε，使得对于所有的$J_0>J_\varepsilon$，$\mathrm{var}\{\overline{V}_{J_0,t}\}<\varepsilon$. 首先，如果我们注意到$\sum_l\tilde{g}_{J_0,l}^2=1/2^{J_0}$(参见表202)，帕塞瓦尔定理告诉我们

$$\int_{-1/2}^{1/2}\widetilde{\mathcal{G}}_{J_0}(f)\,\mathrm{d}f=\sum_{l=0}^{L_{J_0}-1}\tilde{g}_{J_0,l}^2=\frac{1}{2^{J_0}}.$$

如果$S_Y(\cdot)$是有界的，由有限数C界定，我们可以确定，对于所有的$J_0>J_\varepsilon\equiv\log_2(C/\varepsilon)$，

$$\mathrm{var}\{\overline{V}_{J_0,t}\}=\int_{-1/2}^{1/2}\widetilde{\mathcal{G}}_{J_0}(f)S_Y(f)\,\mathrm{d}f\leqslant C\int_{-1/2}^{1/2}\widetilde{\mathcal{G}}_{J_0}(f)\,\mathrm{d}f=\frac{C}{2^{J_0}}<\varepsilon;$$

如果$S_Y(\cdot)$不能被任一有限数C界定，则存在一个常数C_ε使得

$$\int_{S_Y(f)\geqslant C_\varepsilon}S_Y(f)\,\mathrm{d}f<\frac{\varepsilon}{2},$$

其中，现在必须利用勒贝格(Lebesgue)积分. 因为对于在第 4 章所讨论的所有小波，以及所有的 f，$\widetilde{\mathcal{G}}_{J_0}(f)\leqslant 1$，对于所有的 $J_0>J_\varepsilon\equiv\log_2(C_\varepsilon/\varepsilon)+1$，

$$\begin{aligned}\int_{-1/2}^{1/2}\widetilde{\mathcal{G}}_{J_0}(f)S_Y(f)\mathrm{d}f&=\int_{S_Y(f)\geqslant C_\varepsilon}+\int_{S_Y(f)<C_\varepsilon}\widetilde{\mathcal{G}}_{J_0}(f)S_Y(f)\mathrm{d}f\\&\leqslant\int_{S_Y(f)\geqslant C_\varepsilon}S_Y(f)\mathrm{d}f+C_\varepsilon\int_{S_Y(f)<C_\varepsilon}\widetilde{\mathcal{G}}_{J_0}(f)\mathrm{d}f\\&\leqslant\frac{\varepsilon}{2}+C_\varepsilon\int_{-1/2}^{1/2}\widetilde{\mathcal{G}}_{J_0}(f)\mathrm{d}f\leqslant\frac{\varepsilon}{2}+\frac{C_\varepsilon}{2^{J_0}}<\varepsilon,\end{aligned}$$

这建立了式(296d).

[2]正确的是，由于泄漏现象周期图可能是真正谱密度函数的一个严重偏差的估计量(在图 273 中阐明)，并且，用不受边界条件影响的极大重叠离散小波变换的部分构造小波方差的一个无偏估计量是容易的(参见 8.3 节). 符合一个小宽度小波滤波器的小波方差会提供对泄漏形式有贡献的误导或无信息结果也是事实——作为进一步的讨论，参见 8.9 节. 足可以说，小波方差中的泄漏本质上使小波方差和谱密度函数之间的近似联系(297a)不成立. 在周期图中发现和排除泄漏需要利用技术，比如，预先漂白或逐渐减弱，我们仅仅需要增加相应的小波滤波器的宽度 L 来减小小波方差中的泄漏.

[3]给定两个适当的有小波系数$\{\overline{W}_{X,j,t}\}$和$\{\overline{W}_{Y,j,t}\}$的随机过程$\{X_t\}$和$\{Y_t\}$，我们能定义小波协方差

$$\nu_{XY}(\tau_j)\equiv\operatorname{cov}\{\overline{W}_{X,j,t},\overline{W}_{Y,j,t}\},$$

它产生一个$\{X_t\}$和$\{Y_t\}$之间的协方差的基于尺度的分解：

$$\sum_{j=1}^{\infty}\nu_{XY}(\tau_j)=\operatorname{cov}\{X_t,Y_t\}.$$

小波协方差可以标准化以产生一个小波对射变换，即

$$\rho_{XY}(\tau_j)\equiv\frac{\operatorname{cov}\{\overline{W}_{X,j,t},\overline{W}_{Y,j,t}\}}{(\operatorname{var}\{\overline{W}_{X,j,t}\}\operatorname{var}\{\overline{W}_{Y,j,t}\})^{1/2}}=\frac{\nu_{XY}(\tau_j)}{\nu_X(\tau_j)\nu_Y(\tau_j)},$$

其中 $\nu_X^2(\tau_j)$和 $\nu_Y^2(\tau_j)$是$\{X_t\}$和$\{Y_t\}$的小波方差. 用与小波方差类似的方法，小波协方差和对射变换，可以用于研究在尺度对尺度的基础上两个过程是如何联系的(这个方法的细节和应用例子，参见 Hudgins，1992，Hudgins 等，1993；Liu，1994；Lindsay 等，1996；Torrence and Compo，1998；Whitcher，1998；Serroukh and Walden，2000a 和 2000b，Whitcher 等，2000b；及其中的参考文献).

8.2 小波方差的基本性质

假定$\{X_t\}$是一个随机过程，它的 d 阶向后差分

$$Y_t = (1-B)^d X_t \equiv \sum_{k=0}^{d} \binom{d}{k} (-1)^k X_{t-k}$$

是一个具有谱密度函数 $S_Y(\cdot)$和均值 μ_Y 的平稳过程.（详情参见 7.7 节）. 回想

$$\overline{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l}, \quad t = \cdots, -1, 0, 1, \cdots$$

表示用极大重叠离散小波变换小波滤波器$\{\tilde{h}_{j,l}\}$滤波$\{X_t\}$获得的输出. 然后，我们可以断言，如果$\{\tilde{h}_{j,l}\}$是使 $L \geqslant 2d$ 的以 Daubechies 小波滤波器为基础的(如 4.8 节所定义的)，那么$\{\overline{W}_{j,t}\}$是一个平稳过程，其谱密度函数是由

$$S_j(f) = \widetilde{\mathcal{H}}_j^{(D)}(f) S_X(f) \tag{304}$$

定义的，其中$\widetilde{\mathcal{H}}_j^{(D)}(\cdot)$是与$\{\tilde{h}_{j,l}\}$有关的平方增益函数. 为看到此对于 $j=1$ 是正确的，首先指出，因为$\tilde{h}_{1,l} = \tilde{h}_l = h_l/\sqrt{2}$，由式(105b)有

$$\widetilde{\mathcal{H}}_1^{(D)}(f) = \frac{1}{2}\mathcal{H}^{(D)}(f) = \sin^L(\pi f) \sum_{l=0}^{\frac{L}{2}-1} \binom{\frac{L}{2}-1+l}{l} \cos^{2l}(\pi f) = \mathcal{D}^{\frac{L}{2}}(f)\, \widetilde{\mathcal{A}}_L(f),$$

其中

$$\mathcal{D}(f) \equiv 4\sin^2(\pi f) \quad , \quad \widetilde{\mathcal{A}}_L(f) \equiv \frac{1}{2^L} \sum_{l=0}^{\frac{L}{2}-1} \binom{\frac{L}{2}-1+l}{l} \cos^{2l}(\pi f).$$

因此，可以认为$\{h_{1,l}\}$等同于一个两级的级联滤波器(参见 4.8 节的讨论). 级联中的第一个滤波器通过给定$\mathcal{D}^{\frac{L}{2}}(\cdot)$获得一个平方增益函数，这与一个 $L/2$ 阶的向后差分滤波器符合. 令$\{O_t\}$表示第一个滤波器的输出，我们有

$$O_t = (1-B)^{\frac{L}{2}} X_t = (1-B)^{\frac{L}{2}-d}[(1-B)^d X_t] = (1-B)^{\frac{L}{2}-d} Y_t.$$

因为由假设$\{Y_t\}$是一个平稳过程，由此得出，用$\frac{L}{2}-d$ 向后差分滤波器反复滤波$\{Y_t\}$，产生一个过程$\{O_t\}$，它必定是平稳的(参见关于式(268a)与(268b)的讨论). 第二级的滤波器有一个平方增益函数$\widetilde{\mathcal{A}}_L(\cdot)$，由构造，它相应于一个有限宽度的滤波器，因此，用一个平方增益函数为$\widetilde{\mathcal{A}}_L(\cdot)$的滤波器滤波$\{O_t\}$得到的输出$\{\overline{W}_{1,t}\}$是一个平稳过程，它的谱密度函数由 $S_1(f) \equiv \widetilde{\mathcal{H}}_1^{(D)}(f) S_X(f)$确定，图 305 显示的是一个滤波器级联的流程图，它是从$\{X_t\}$产生$\{\overline{W}_{1,t}\}$，而$\{Y_t\}$和$\{O_t\}$形成的中间过程.

练习[304] 建立 $j \geqslant 2$ 时的结论. 提示：$\sin^2(2f) = 4\sin^2(f)\cos^2(f)$. ◁

$$X_t \longrightarrow \boxed{\mathcal{D}^d(\cdot)} \longrightarrow Y_t \longrightarrow \boxed{\mathcal{D}^{\frac{L}{2}-d}(\cdot)} \longrightarrow O_t \longrightarrow \boxed{\widetilde{\mathcal{A}}_L(\cdot)} \longrightarrow \overline{W}_{1,t}$$

图 305 流程图描述了用 Daubechies 极大重叠离散小波变换小波滤波器$\{\tilde{h}_l\}$滤波$\{X_t\}$得到$\overline{W}_{1,t}$. 实际的滤波器可以分解成一个含有三个滤波器的级联(如上显示的通过它们的平方增益函数$\mathcal{D}^d(\cdot)$、$\mathcal{D}^{\frac{L}{2}-d}(\cdot)$和$\widetilde{\mathcal{A}}_L(\cdot)$). 由假设，$\{X_t\}$的 d 阶向后差分是一个平稳过程$\{Y_t\}$，同时小波滤波器具有宽 $L\geqslant 2d$，因此包含 $L/2$ 向后差分滤波器$\mathcal{D}(\cdot)$. 当 $L=2d$ 时，过程$\{Y_t\}$和$\{O_t\}$是相等的(滤波器$\mathcal{D}^0(\cdot)$是一个"什么也不做"的滤波器)；当 $L>2d$ 时，通过进行$\frac{L}{2}-d$ 次逐次向后差分$\{Y_t\}$得到$\{O_t\}$. 过程$\{O_t\}$通过光滑滤波器$\widetilde{\mathcal{A}}_L(\cdot)$的作用产生第一层小波系数过程$\{\overline{W}_{1,t}\}$

练习[305] 证明：一方面如果 $L>2d$ 或 $L=2d$，且 $\mu_Y=0$，那么 $E\{\overline{W}_{j,t}\}=0$；另一方面，如果 $\mu_Y\neq 0$，$L=2d$，那么 $E\{\overline{W}_{j,t}\}\neq 0$. ◁

由于平稳过程的方差等于它的谱密度函数积分，并且由于$\{\overline{W}_{j,t}\}$的方差是小波方差 $\nu_X^2(\tau_j)$，则有

$$\nu_X^2(\tau_j)=\int_{-1/2}^{1/2} S_j(f)\,\mathrm{d}f=\int_{-1/2}^{1/2}\widetilde{\mathcal{H}}_j^{(\mathrm{D})}(f)S_X(f)\,\mathrm{d}f, \tag{305}$$

只要小波滤波器的宽度 L 至少为 $2d$(为了将$\{X_t\}$转化成一个平稳过程，我们必须应用两次向后差分). 特别地，对于 $j=1$ 有

$$\nu_X^2(\tau_1)=\int_{-1/2}^{1/2}\widetilde{\mathcal{H}}_1^{(\mathrm{D})}(f)S_X(f)\,\mathrm{d}f=2^{L-2d}\int_{-1/2}^{1/2}\widetilde{\mathcal{A}}_L(f)\sin^{L-2d}(\pi f)S_Y(f)\,\mathrm{d}f.$$

条件 $L\geqslant 2d$ 保证了被积函数中的正弦项界为 1，因而此积分是有限的.

注意当小波滤波器的宽足够大时，无论在平稳过程还是有着平稳向后差分的非平稳过程，小波方差都会得到很好的定义. 对于平稳过程，任一 L 满足：我们有 $E\{\overline{W}_{j,t}\}=0$；且在所有尺度上的小波方差的和等于过程的方差(式(296d)). 对于具有 d 阶平稳向后差分的非平稳过程，我们需要 $L\geqslant 2d$. 我们可以要求更强的条件 $L>2d$ 以得到 $E\{\overline{W}_{j,t}\}=0$；而小波方差的和趋于无穷. 典型地，统计就是设计平稳过程的工作(例如，不同的谱密度函数估计)对于平稳的向后差分的非平稳过程，需要特殊的方法，以便研究它们的特征(例如，偏离性和方差). 小波的一个优点在于理论应用对于两种过程同样适合.

8.2 节的评论与扩展

[1]这里，为了简化本章更进一步的讨论，我们假设$\{\tilde{h}_{j,l}\}$是由一个宽为 L 的 Daubechies 小波滤波器构造的. 因为此滤波器包含了一个嵌入 $L/2$ 阶向后差分的滤波器. 我们必须利用条件 $L\geqslant 2d$ 去保证此滤波器可以恰当地应用于有着平稳 d 阶向后差分的过程. 如 4.9 节所讨论的，coiflet 滤波器也包含一个嵌入向后差分滤波器，因此它们当然可以代替 Daubechies 滤波器. 然而宽为 L 的滤波器和差分滤波器的阶数之间的关系是 $L/3$，本节的结论在 $L\geqslant 3d$ 的条件下需要重新表示(它相当于简单地分别替代所有的 $2d$ 和 $L/2$，而用 $3d$ 和 $L/3$).

8.3 小波方差的估计

现在假设我们给定一个时间序列，它能够认为是过程$\{X_t\}$的一部分 X_0，…，X_{N-1}的实现，过程$\{X_t\}$的 d 阶向后差分形成一个平稳过程，我们想利用宽度为 $L\geqslant 2d$ 的 Daubechies 小波滤波器去估计小波方差 $\nu_X^2(\tau_j)$，其中我们假定 L 足够大，以使 $E\{\overline{W}_{j,t}\}=0$(参见练习[305]). 在此假设下我们有

$$\nu_X^2(\tau_j) = \mathrm{var}\{\overline{W}_{j,t}\} = E\{(\overline{W}_{j,t} - E\{\overline{W}_{j,t}\})^2\} = E\{\overline{W}_{j,t}^2\},$$

因此可以将估计建立在平方过程$\{\overline{W}_{j,t}^2\}$上.

如果注意不包括包含循环假设的极大重叠离散小波变换系数，我们可以基于 X_0，…，X_{N-1}的极大重叠离散小波变换得到 $\nu_X^2(\tau_j)$的一个无偏估计量. 为了把这一点解释清楚，回忆第 5 章给出的极大重叠离散小波变换

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N}, \quad t = 0,1,\cdots,N-1, \tag{306a}$$

而在本章中，我们定义

$$\overline{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l}, \qquad t = \cdots,-1,0,1,\cdots.$$

现在，对于这些指标 t，$\widetilde{W}_{j,t}=\overline{W}_{j,t}$. 使 $\widetilde{W}_{j,t}$的结构保持不变，如果我们想消除模运算，只需$t\geqslant L_j-1$，此结论都是成立的. 所以，如果 $N-L_j\geqslant 0$，我们可以基于极大重叠离散小波变换建立一个 $\nu_X^2(\tau_j)$的估计量

$$\hat{\nu}_X^2(\tau_j) \equiv \frac{1}{M_j}\sum_{t=L_j-1}^{N-1} \widetilde{W}_{j,t}^2 = \frac{1}{M_j}\sum_{t=L_j-1}^{N-1} \overline{W}_{j,t}^2, \tag{306b}$$

其中 $M_j\equiv N-L_j+1$. 因为 $E\{\hat{\nu}_X^2(\tau_j)\}=\nu_X^2(\tau_j)$，可以将$\hat{\nu}_X^2(\tau_j)$看作是小波方差的无偏极大重叠离散小波变换估计量.

相反，我们将

$$\tilde{\nu}_X^2(\tau_j) \equiv \frac{1}{N}\sum_{t=0}^{N-1} \widetilde{W}_{j,t}^2 = \frac{1}{N}\left(\sum_{t=0}^{L_j-2} \widetilde{W}_{j,t}^2 + \sum_{t=L_j-1}^{N-1} \overline{W}_{j,t}^2\right) \tag{306c}$$

看作是小波方差的有偏极大重叠离散小波变换估计量(即使有些过程——例如，白噪声过程——使得 $E\{\tilde{\nu}_X^2(\tau_j)\}=\nu_X^2(\tau_j)$)．由于极大重叠离散小波变换的方差分析特性(参见 5.4 节)，这个有偏估计量可以引起一个样本方差的精确分解这当然是吸引人的结论(参见以下评论与扩展的第[4]条)．注意随着 $N\to\infty$，两个估计量的项数比，即$\frac{M_j}{N}$，趋于 1(参见练习[8.5])．下面，我们将集中论述无偏估计量$\hat{\nu}_X^2(\tau_j)$.

由于估计量$\hat{\nu}_X^2(\tau_j)$是一个随机变量(RV)，研究它接近于 $\nu_X^2(\tau_j)$的程度是很有意义的．为了讲清这个问题，我们假设$\{\overline{W}_{j,t}\}$是一个有零均值和谱密度函数 $S_j(\cdot)$的高斯平稳过程，由$\{\overline{W}_{j,t}\}$的任何 N 个随机过程的聚集服从多元正态(或高斯)分布．下述基本结论成立(Percival，1983，1995)：如果几乎处处有 $S_j(f)>0$，并且如果

$$A_j \equiv \int_{-1/2}^{1/2} S_j^2(f)\mathrm{d}f < \infty, \tag{307a}$$

那么式(306b)的估计量$\hat{\nu}_X^2(\tau_j)$是具有均值 $\nu_X^2(\tau_j)$和大样本方差为$\frac{2A_j}{M_j}$的渐进正态分布(在实际中，术语几乎处处所说的是对于所有 f，$S_j(f)\geqslant 0$，则 $S_j(\cdot)$可以在一个孤立的单个频率上为零，但不可能在一个频率区间恒为零)．因此，为对于大的 $M_j=N-L_j+1$ 的一个好的逼近，我们有

$$\frac{M_j^{1/2}(\hat{\nu}_X^2(\tau_j)-\nu_X^2(\tau_j))}{(2A_j)^{1/2}} \stackrel{\mathrm{d}}{=} \mathcal{N}(0,1) \tag{307b}$$

(参见式(257a))．

为了说明(307a)中的平方可积条件，我们注意到，因为$\{\overline{W}_{j,t}\}$是一个平稳过程，它有一个自协方差序列$\{s_{j,\tau}\}$．如果自协方差序列满足式(267a)中的平方可和条件，则 $\{s_{j,\tau}\}\leftrightarrow S_j(\cdot)$；即自协方差序列和谱密度函数构成一个傅里叶变换对．由帕塞瓦尔定理(参见式 35c)可得

$$\sum_{\tau=-\infty}^{\infty} s_{j,\tau}^2 = \int_{-1/2}^{1/2} S_j^2(f)\mathrm{d}f.$$

因此，只要$\{\overline{W}_{j,t}\}$的自协方差序列衰减的足够快以致于它可平方求和，则平方可积条件成立．事实上，如果 $S_j(\cdot)$不平方可积，由于在 $f=0$ 的奇异性，它可以通过增加小波滤波器的宽度得以改善．

为了得到上面的大样本结论是不是中样本的一个好的逼近，假设$\{X_t\}$要么是一个平稳分形差分(FD)过程，要么是一个非平稳分形过程，其向后差分阶数 $d=1$ 和 $d=2$ 形成一个平稳分形差分过程(详情参见 7.6 和 7.7 节)．集合这些过程是一个幂规律指数 α 取遍 $-5<\alpha\leqslant 1$ 的 $1/f$ 型过程．仿真研究(Percival，1995)显示了式(307b)中的大样本方差是样本方差$\hat{\nu}_X^2(\tau_1)$的准确衡量，2%范围内，N 取得使 $M_1=128$ 对于

[1]具有 $\delta=-\frac{1}{2}$，$-\frac{1}{4}$，$-\frac{1}{8}$，0，$\frac{1}{8}$，$\frac{1}{4}$，$\frac{1}{2}$，1 和 $\frac{3}{2}$ 的分形差分过程；

[2]宽为 $L=2$，4，6 和 8 的 Daubechies 小波滤波器

所有的可能组合(除了 $L=2$ 和 $\delta=\frac{3}{2}$ 组合外，在这种组合下，(307a)所要求的平方可积条件失效)，因此，大样本理论是相应的小样本理论的近似.

8.3 节的评论与扩展

[1]我们也可以用离散小波变换代替极大重叠离散小波变换估计小波方差. 在假设 $\{X_t\}$ 的样本大小 N 能用 2^j 整除的条件下，我们首先计算 j 层小波系数 $W_{j,t}$，$t=0$，…，N_j-1，和以前一样，$N_j\equiv\frac{N}{2^j}$. 我们能够形成 $\nu_X^2(\tau_j)$ 的一个无偏估计，而用通过去除边界系数(这些明确了用循环假设)和平均剩余系数的平方(重新规范化后，考虑离散小波变换和极大重叠离散小波变换的系数关系，即 $W_{j,t}/2^{j/2}=\widetilde{W}_{j,2^j(t+1)-1}$). 如 4.11 节所讨论的，边界系数的数目是 $\min\{L'_j,N_j\}$，其中 L'_j 由式(146a)确定，不会超过 $L-2$. 在假设 $N_j>L'_j$ 下，基于离散小波变换上的无偏估计量由下式给出：

$$\hat{\nu}_X^2(\tau_j)\equiv\frac{1}{(N_j-L'_j)2^j}\sum_{t=L'_j}^{N_j-1}W_{j,t}^2=\frac{1}{N-2^jL'_j}\sum_{t=L'_j}^{N_j-1}W_{j,t}^2. \tag{308a}$$

我们也可以基于离散小波变换定义一个有偏估计量

$$\tilde{\nu}_X^2(\tau_j)\equiv\frac{1}{N_j2^j}\sum_{t=0}^{N_j-1}W_{j,t}^2=\frac{1}{N}\sum_{t=0}^{N_j-1}W_{j,t}^2. \tag{308b}$$

如果样本大小 N 不能用 2^j 整除，我们仍然可以得到基于离散小波变换上的一个小波方差的无偏估计量，如下. 令 $M>N$ 为被 2^j 整除的 2 的任意次幂，构造一个新的长为 M 的时间序列 $\{X'_t\}$，通过对原序列 $\{X_t\}$ 添加 $M-N$ 个 0：

$$X'_t\equiv\begin{cases}X_t, & 0\leqslant t\leqslant N-1;\\ 0, & N\leqslant t\leqslant M-1.\end{cases} \tag{308c}$$

令 $W'_{j,t}$ 表示 $\{X'_t\}$ 的 j 层小波系数，它很容易计算，因为 M 能被 2^j 整除. 练习[8.7]的要点就是去证明基于用 $2^{j/2}$ 的划分，这些系数的子序列，即

$$W'_{j,L'_j},W'_{j,L'_j+1},\cdots,W'_{j,\lfloor\frac{N}{2^j}-1\rfloor}, \tag{308d}$$

等于从无穷序列 $\{\widetilde{W}_{j,t}\}$ 上选出的随机变量，因而需要的无偏估计量为

$$\hat{\nu}_X^2(\tau_j)\equiv\frac{1}{\left(\left\lfloor\frac{N}{2^j}-1\right\rfloor-L'_j+1\right)2^j}\sum_{t=L'_j}^{\lfloor\frac{N}{2^j}-1\rfloor}[W'_{j,t}]^2=\frac{1}{M'_j2^j}\sum_{t=L'_j}^{\lfloor\frac{N}{2^j}-1\rfloor}\lfloor W'_{j,t}\rfloor^2,$$

其中 $M_j' \equiv \left\lfloor \frac{N}{2^j} - 1 \right\rfloor - L_j' + 1$(注意到这与以前的当 N 被 2^j 整除时的结论相符合).

为了得到无偏极大重叠离散小波变换和离散小波变换估计量相对指标的一些信息，让我们专门研究 $\nu_X^2(\tau_1)$，即单位尺度方差. 在假设 $\{\overline{W}_{1,t}\}$ 的自协方差序列 $\{s_{1,\tau}\}$ 是平方可求和的，它对谱密度函数满足 $S_1(f)>0$ 几乎处处成立的条件下，$\hat{\hat{\nu}}_X^2(\tau_1)$ 服从渐进正态分布，它的均值为 $\nu_X^2(\tau_j)$，大样本方差为

$$\frac{1}{2M_1'}\int_{-1/2}^{1/2}\left(S_1\left(\frac{f}{2}\right)+S_1\left(\frac{f}{2}+\frac{1}{2}\right)\right)^2 \mathrm{d}f$$

(Percival，1995). 与这个结论相平行的极大重叠离散小波变换上面被积函数是子样本随机过程

$$2^{1/2}\overline{W}_{1,2t+1} = \sum_{l=0}^{L-1} h_l X_{2t+1-l}, \quad t = \cdots, -1, 0, 1, \cdots$$

的谱密度函数的平方(参见练习[8.8]). 现在我们将前者和后者的渐近相对效率 $\hat{\hat{\nu}}_X^2(\tau_1)$ 与 $\hat{\nu}_X^2(\tau_1)$ 进行比较：

$$e(\hat{\hat{\nu}}_X^2(\tau_1), \hat{\nu}_X^2(\tau_1)) \equiv \lim_{N\to\infty}\frac{\operatorname{var}\{\hat{\nu}_X^2(\tau_1)\}}{\operatorname{var}\{\hat{\hat{\nu}}_X^2(\tau_1)\}} = \frac{\int_{-1/2}^{1/2} S_1^2(f)\mathrm{d}f}{\int_{-1/2}^{1/2} S_1^2(f)\mathrm{d}f + \int_{-1/2}^{1/2} S_1\left(\frac{f}{2}\right)S_1\left(\frac{f}{2}+\frac{1}{2}\right)\mathrm{d}f} \tag{309}$$

(练习[8.9]验证了以上结论). 因为任意谱密度函数是非负函数，显然 $e(\hat{\hat{\nu}}_X^2(\tau_1),\ \hat{\nu}_X^2(\tau_1)) \leqslant 1$，它意味着，给同样大的数据量，离散小波变换的估计量没有基于极大重叠离散小波变换上的估计量有效，感觉上前者的方差没有后者的小，然而，就下面的讨论似乎渐近相对效率应该十分接近于 1，因为我们希望第二个积分接近于 0，回忆

$$S_1(f) = \widetilde{\mathcal{H}}^{(\mathrm{D})}(f)S_X(f) \approx \begin{cases} S_X(f), & \frac{1}{4} \leqslant |f| \leqslant \frac{1}{2}; \\ 0, & 0 \leqslant |f| \leqslant \frac{1}{4}, \end{cases}$$

其中近似值对于小波滤波器类似于一个半带高通滤波器是有效的. 因为 $S_1(\cdot)$ 是有着单位周期的周期函数，由此得出

$$S_1\left(f+\frac{1}{2}\right) \approx \begin{cases} 0, & \frac{1}{4} \leqslant |f| \leqslant \frac{1}{2}; \\ S_X\left(f+\frac{1}{2}\right), & 0 \leqslant |f| \leqslant \frac{1}{4}. \end{cases}$$

这对于所有的 f 产生近似 $S_1\left(\frac{f}{2}\right)S_1\left(\frac{f}{2}+\frac{1}{2}\right) \approx 0$，又暗示了渐近相对效率应该接近于 1(如果小波滤波器是一个理想的高通滤波器就是 1). 通过这个结论，使用(计算快)基于离散小波变换的估计量代替极大重叠离散小波变换估计量的误差不会太大.

然而，如果我们融合具体的不同宽度的小波滤波器对某个特殊的过程 $\{X_t\}$ 计算渐近相对效率，就会发现，事实上，离散小波变换估计量效率相当低，更糟糕

表　310

δ	$L=2$	$L=4$	$L=6$	$L=8$
$-1/2$	0.85	0.89	0.91	0.92
$-1/4$	0.81	0.86	0.89	0.90
$-1/8$	0.78	0.84	0.87	0.89
0	0.75	0.82	0.85	0.87
1/8	0.72	0.80	0.83	0.86
1/4	0.68	0.77	0.81	0.84
1/2	0.61	0.72	0.77	0.80
1	0.50	0.61	0.67	0.71
3/2	—	0.52	0.58	0.62

注：离散小波变换估计$\hat{\hat{\nu}}_X^2(\tau_1)$关于基于极大离散小波变换的估计$\hat{\nu}_X^2(\tau_1)$对于不同的分形差分过程和Daubechies小波滤波器的组合的渐近相对效率(AREs). 注意一个小于1的渐近相对效率意味着$\hat{\nu}_X^2(\tau_1)$有着比$\hat{\hat{\nu}}_X^2(\tau_1)$小的大样本方差.

的是，它的大样本方差是极大重叠离散小波变换估计量的两倍. 这个事实在表310中说明. 这个表给出了δ从$-1/2$到$3/2$，$L=2$，4，6和8的分形差分过程的渐近相对效率. 因为一个分形差分过程有一个谱密度函数与$|\sin(\pi f)|^{-2\delta}$成比例，我们看到，将L固定，渐近相对效率随着谱密度函数的低频部分减小越来越明显. 还注意到，将δ固定，渐近相对效率随着L的增加而增加. 这是合理的，因为渐近相对效率小于1是由于小波滤波器是作为一个理想高通滤波器，小波滤波器有其不完全性，随着L的增加，小波滤波器成为此滤波器的一个好的近似. 因此我们通过用极大重叠离散小波变换估计量代替离散小波变换估计量得到两个因子的大样本方差的简化.

[2]通过极大重叠离散小波变换用塔式算法很容易得到无偏估计量$\hat{\nu}_X^2(\tau_j)$；但是由于为了构造这个估计量抛弃了所有的边界系数，相对于真正需要的来说，我们会计算更多的极大重叠离散小波变换系数. 采用极大重叠离散小波变换塔式算法是可能的，因此只需计算$\hat{\nu}_X^2(\tau_j)$要用的系数，详情参见Percival and Guttorp(1994).

[3]假设$\{X_t: t=0, 1, \cdots, N-1\}$是一个长为$N$的幂规律过程的一个部分，其中$-2\leqslant\alpha\leqslant0$. 最近，Greenhall等(1999)证明了，如果$\{X_t\}$就像式(140)通过与反演序列$\{X_t, t=N-1, N-2, \cdots, 0\}$合并，扩展成一个长为$2N$的序列，长为$2N$的序列的哈尔小波方差的有偏极大重叠离散小波变换估计量在尺度上接近于$\frac{N}{2}$时，比原序列的无偏估计量有着明显的均值平方误差. 这个结果提示了通过构造一个特别循环量处理循环假设对于长的记忆过程有很多理论上的证明，并且有偏极大重叠离散小波变换小波方差估计量对于除哈尔小波外可能有很多重要的优点(这两个问题现在普遍在研究).

[4]最近Serroukh等(2000)得到了广泛过程的极大重叠离散小波变换小波方差估计量的渐进分布，它并不必要是高斯的或线性的，并且是(i)平稳的，或(ii)不平稳但有平稳的d次差分，或(iii)全局不平稳但局部平稳. 他们发现，在这些

(和附加)条件下，估计量$\hat{\nu}_X^2(\tau_j)$又服从均值为$\nu_X^2(\tau_j)$的渐近高斯分布，但是现在通过$S_{W_j^2}(0)/M_j$给出大样本方差，其中$S_{W_j^2}(\cdot)$是过程$\{\overline{W}_{j,t}^2\}$的谱密度函数(实际上，估计$S_{W_j^2}(\cdot)$在$f=0$处的一个好的方法，是通过一个多锥估计量)．在小尺度下，用于非高斯时间序列的极大重叠离散小波变换估计量的方差的估计比在高斯假设条件下计算出的大得多．平稳过程$\{\overline{W}_{j,t}\}$的高斯假设导致一个大样本结论(式(307b))．

8.4 小波方差的置信区间

式(307b)的大样本结论主要用于基于$\hat{\nu}_X^2(\tau_j)$计算$\nu_X^2(\tau_j)$的置信区间，为了做到这一点，令$\Phi^{-1}(p)$表示标准高斯分布的$p\times100\%$百分比(参见式(257b))．高斯分布的对称特性意味着对于$0\leqslant p\leqslant 1/2$，

$$\mathbf{P}[-\Phi^{-1}(1-p)\leqslant Z\leqslant\Phi^{-1}(1-p)]=1-2p.$$

对于大的N，式(307b)告诉我们近似地

$$\mathbf{P}\left[-\Phi^{-1}(1-p)\leqslant\frac{M_j^{1/2}(\hat{\nu}_X^2(\tau_j)-\nu_X^2(\tau_j))}{(2A_j)^{1/2}}\leqslant\Phi^{-1}(1-p)\right]=1-2p.$$

现在括号中的事件发生当且仅当条件

$$-\Phi^{-1}(1-p)\leqslant\frac{M_j^{1/2}(\hat{\nu}_X^2(\tau_j)-\nu_X^2(\tau_j))}{(2A_j)^{1/2}}$$

和

$$\frac{M_j^{1/2}(\hat{\nu}_X^2(\tau_j)-\nu_X^2(\tau_j))}{(2A_j)^{1/2}}\leqslant\Phi^{-1}(1-p)$$

同时发生；同样，这两个事件又分别等价于

$$\nu_X^2(\tau_j)\leqslant\hat{\nu}_X^2(\tau_j)+\Phi^{-1}(1-p)\left(\frac{2A_j}{M_j}\right)^{1/2}$$

和

$$\hat{\nu}_X^2(\tau_j)-\Phi^{-1}(1-p)\left(\frac{2A_j}{M_j}\right)^{1/2}\leqslant\nu_X^2(\tau_j).$$

把这两事件结合起来等价于单一事件

$$\hat{\nu}_X^2(\tau_j)-\Phi^{-1}(1-p)\left(\frac{2A_j}{M_j}\right)^{1/2}\leqslant\nu_X^2(\tau_j)\leqslant\hat{\nu}_X^2(\tau_j)+\Phi^{-1}(1-p)\left(\frac{2A_j}{M_j}\right)^{1/2},$$

它发生的概率为$1-2p$．因此，我们可以说，以概率$1-2p$，随机区间

$$\left[\hat{\nu}_X^2(\tau_j)-\Phi^{-1}(1-p)\left(\frac{2A_j}{M_j}\right)^{1/2},\hat{\nu}_X^2(\tau_j)+\Phi^{-1}(1-p)\left(\frac{2A_j}{M_j}\right)^{1/2}\right]\tag{311}$$

包含 $\nu_X^2(\tau_j)$ 真实值和为 $\nu_X^2(\tau_j)$ 构造了一个 $100\times(1-2p)\%$ 的置信区间.

在实际中为了使用式(311)，我们必须估计 $S_j^2(\cdot)$ 的积分 A_j. 由于 A_j 大部分由 $S_j(\cdot)$ 的大值决定，并且，由于周期图可以在相对小偏差的情况下估计谱密度函数的大值，我们可以使用周期图作为 $S_j(\cdot)$ 的评估值：

$$\hat{S}_j^{(p)}(f) \equiv \frac{1}{M_j}\left|\sum_{t=L_j-1}^{N-1}\widetilde{W}_{j,t}e^{-i2\pi ft}\right|^2.$$

由式(269b)知，对于 $0<|f|<1/2$ 和大的 N，比

$$\frac{2\hat{S}_j^{(p)}(f)}{S_j(f)} \overset{d}{=} \chi_2^2,$$

其中 χ_2^2 是一个有 2 阶自由度的 χ^2 随机过程(回想在 7.2 节中 χ^2 分布的性质).

练习[312] 用大样本结果讨论

$$\hat{A}_j \equiv \frac{1}{2}\int_{-1/2}^{1/2}[\hat{S}_j^{(p)}(f)]^2 df \tag{312a}$$

对于大的 M_j 是 A_j 的近似无偏估计. ◁

由式(269a)可得 $\{\hat{s}_{j,\tau}^{(p)}\}\longleftrightarrow\hat{S}_j^{(p)}(\cdot)$，其中 $\hat{s}_{j,\tau}^{(p)}$ 是自协方差序列通常的有偏估计，即

$$\hat{s}_{j,\tau}^{(p)} \equiv \frac{1}{M_j}\sum_{t=L_j-1}^{N-1-|\tau|}\widetilde{W}_{j,t}\widetilde{W}_{j,t+|\tau|}, \quad 0\leqslant|\tau|\leqslant M_j-1, \tag{312b}$$

并且，对于 $|\tau|\geqslant M_j$ 有 $\hat{s}_{j,\tau}^{(p)}\equiv 0$. 引用帕塞瓦尔定理(式(35c))，我们有

$$\int_{-1/2}^{1/2}[\hat{S}_j^{(p)}(f)]^2 df = \sum_{\tau=-(M_j-1)}^{M_j-1}(\hat{s}_{j,\tau}^{(p)})^2.$$

因而，我们得到简便的计算公式

$$\hat{A}_j = \frac{(\hat{s}_{j,0}^{(p)})^2}{2} + \sum_{\tau=1}^{M_j-1}(\hat{s}_{j,\tau}^{(p)})^2 = \frac{\hat{\nu}_X^4(\tau_j)}{2} + \sum_{\tau=1}^{M_j-1}(\hat{s}_{j,\tau}^{(p)})^2.$$

在式(311)中用 $\hat{A}_j$ 代换 A_j 导致了对于 $\nu_X^2(\tau_j)$ 的一个近似 $100(1-2p)\%$ 置信区间(以有效的假设为条件，估计量 $\hat{A}_j$ 接近于 A_j).

式(311)给出典型的置信区间的一个难点是不能防止置信下界是负的，它是一个冗余，因为我们知道 $\nu_X^2(\tau_j)$ 必定非负，而且因为我们经常想在对数尺度下为

置信区间绘图. 现在$\hat{\nu}_X^2(\tau_j)$与 M_j 个零均值高斯随机变量的平方和成比例，每个高斯随机变量有着相同的方差. 如果这些随机变量均不相关，式(263a)指出$\hat{\nu}_X^2(\tau_j)$能够重新规范化以服从 M_j 自由度的 χ- 分布，随后由它可以得到一个严格正的置信区间(它的形式与下面的式(313c)类似). 我们可以逼近用尺度化的 χ- 分布得到有着零均值和相同方差的高斯随机过程相关联的的平方和的近似分布，这个 χ- 分布有着足以说明随机变量间关联的自由度(Priestley, 1981, p. 466). 在这个途径下，我们使用逼近

$$\frac{\eta\hat{\nu}_X^2(\tau_j)}{\nu_X^2(\tau_j)} \stackrel{d}{=} \chi_\eta^2, \tag{313a}$$

其中 η 是著名的"等价自由度"(EDOF)，且确定如下.

练习[313a] 在假设$\hat{\nu}_X^2(\tau_j)$如 $a\chi_\eta^2$ 分布之下，通过解两个方程

$$E\{\hat{\nu}_X^2(\tau_j)\} = E\{a\chi_\eta^2\}$$

和

$$\operatorname{var}\{\hat{\nu}_X^2(\tau_j)\} = \operatorname{var}\{a\chi_\eta^2\}$$

确定 a 和 η(等价自由度). 特别地，证明

$$\eta = \frac{2(E\{\hat{\nu}_X^2(\tau_j)\})^2}{\operatorname{var}\{\hat{\nu}_X^2(\tau_j)\}}.$$

利用大样本$\hat{\nu}_X^2(\tau_j)$的均值和方差的近似方法证明等价自由度由下式近似地给出：

$$\eta_1 \equiv \frac{M_j\nu_X^4(\tau_j)}{A_j}. \tag{313b}$$

◁

式(313a)中的假设分布事实上当 $M_j=1$ 时很精确，并且当 $N\to\infty$时渐近正确(因为当自由度趋于正无穷时 χ^2 分布趋于正态分布).

练习[313b] 在假设式(313a)成立的条件下，利用导出式(311)的类似论证证明 $\nu_X^2(\tau_j)$近似的 $100(1-2p)\%$置信区间由

$$\left[\frac{\eta\hat{\nu}_X^2(\tau_j)}{Q_\eta(1-p)}, \frac{\eta\hat{\nu}_X^2(\tau_j)}{Q_\eta(p)}\right] \tag{313c}$$

给出，其中 $Q_\eta(p)$是 χ_η^2 分布的 $p\times100\%$百分比点，即 $\mathbf{P}[\chi_\eta^2\leqslant Q_\eta(p)]=p$. ◁

在实际应用中，计算上面置信区间需要的自由度 η 由下式估计：

$$\hat{\eta}_1 \equiv \frac{M_j\hat{\nu}_X^4(\tau_j)}{\hat{A}_j}. \tag{313d}$$

另外一种获得 $\nu_X^2(\tau_j)$置信区间的方法是假设 $S_X(\cdot)$，并且因此 $S_j(\cdot)$用一常数相乘，即我们假定 $S_j(f)=aC_j(f)$，其中 $C_j(\cdot)$是一个已知函数，a 是一个未知常数．这个假设，例如，在 Greenhall(1991)为获得 Allan 方差的置信区间中使用过(参见 8.6 节)．由于周期图与 $\widetilde{W}_{j,t}$的离散傅里叶变换的平方模成比例，应用帕塞瓦尔定理，得到

$$\sum_{t=L_j-1}^{N-1}\widetilde{W}_{j,t}^2=\sum_{k=0}^{M_j-1}\hat{S}_j^{(p)}(f_k),$$

其中 $f_k\equiv k/M_j$，因此我们有

$$\hat{\nu}_X^2(\tau_j)=\frac{1}{M_j}\sum_{k=0}^{M_j-1}\hat{S}_j^{(p)}(f_k)\approx\frac{2}{M_j}\sum_{k=1}^{\lfloor(M_j-1)/2\rfloor}\hat{S}_j^{(p)}(f_k),\tag{314a}$$

其中$\lfloor(M_j-1)/2\rfloor$是小于等于$(M_j-1)/2$ 的最大整数．上面的近似仅仅说明 $\hat{S}_j^{(p)}(0)/M_j$($\widetilde{W}_{j,t}$的样本均值的平方)和$\hat{S}_j^{(p)}\left(\frac{1}{2}\right)/M_j$ 是可以忽略的.

练习[314a] 在通常的大样本近似下，对于 $0<f_k<\frac{1}{2}$，

$$\frac{2\hat{S}_j^{(p)}(f_k)}{S_j(f_k)}\stackrel{d}{=}\chi_2^2,$$

且式(314a)右边求和中的随机变量是相互独立的，使用等价自由度讨论要求，近似地有

$$\frac{\eta_2\hat{\nu}_X^2(\tau_j)}{\nu_X^2(\tau_j)}\stackrel{d}{=}\chi_{\eta_2}^2,$$

其中

$$\eta_2\equiv\frac{2\left(\sum_{k=1}^{\lfloor(M_j-1)/2\rfloor}C_j(f_k)\right)^2}{\sum_{k=1}^{\lfloor(M_j-1)/2\rfloor}C_j^2(f_k)}.\tag{314b}$$

◁

$\nu_X^2(\tau_j)$的一个近似 $100(1-2p)\%$置信区间由式(313c)用 η_2代替 η 给出.

一个更进步的简化是回想极大重叠离散小波变换小波滤波器$\{\tilde{h}_{j,l}\}$可以认为是近似带通滤波器，它的通带是 $1/2^{j+1}<|f|\leqslant1/2^j$．这个事实暗示在一些实际问题中假设 $S_j(\cdot)$是带限的，并且恰好覆盖它的标称带宽是合理的(这个假设的有效性通过检查 $S_j(\cdot)$的估计来评价).

练习[314b] 证明等价自由度理论现在导致

$$\eta_3=\max\{M_j/2^j,1\}.\tag{314c}$$

◁

如果样本足够大(在 Percival，1995 的模拟试验报告中显示 $M_j=128$ 通常已足够)，用式(313d)的 η_1 来估计基于具有 η 的式(313c)的置信区间是正确的，我们推荐这种方法．对于小样本，这个方法在有些情况下可以产生优化置信区间，在这些情况下若选择(314b)中的 η_2 或者(314c)中的 η_3，则置信区间都是有用的检验，并且如果相对于选择 η_1 置信区间更宽，我们应该选择 η_2 和 η_3．使用 η_2 要求对 $S_j(\cdot)$的形状有合理的猜测，如果这个猜测不是有效的，则等价自由度应该基于 η_3．

8.4 节的评论与扩展

[1]在实际中，通过周期图估计 $S_j(\cdot)$需要形成的$\{\widetilde{W}_{j,t}: t=L_j-1, \cdots, N-1\}$的离散傅里叶变换，这将是有问题的，如果序列长 $M_j=N-L_j+1$ 非常大．为了避开这个难点，我们可以用 $S_j(\cdot)$的替换的估计量，它把此序列分成不相重叠的块，每一块都包含 N_S 个邻接的小波系数，然后我们总共有 $N_B=\lfloor M_j/N_S\rfloor$块．如果为每块计算一个周期图，平均 N_B 个周期图，我们会获得一个谱密度函数估计，它的形式是

$$\hat{S}_j^{(\mathrm{sa})}(f)\equiv\frac{1}{N_B}\sum_{n=0}^{N_B-1}\left(\frac{1}{N_S}\left|\sum_{t=0}^{N_S-1}\widetilde{W}_{j,L_j-1+nN_S+t}\mathrm{e}^{-\mathrm{i}2\pi ft}\right|^2\right).$$

这个 $S_j(\cdot)$的分段平均估计可以认为是 Welch 重叠分段平均(WOSA)谱密度函数估计(Welch，1967)的一个特殊情形．关于 Welch 重叠分段平均的细节(包括选择一个恰当的块大小 N_S 的指导)，例如，参见 Percival and Walden(1993)，6.17 节．

现在我们可以用$\hat{S}_j^{(\mathrm{sa})}(\cdot)$代替$\hat{S}_j^{(\mathrm{p})}(\cdot)$来估计 A_j，但是需要如下面练习显示的那样来稍微调整式(312a)．

练习[315] 在一些合理的条件下，统计理论提出 $2N_B\hat{S}_j^{(\mathrm{sa})}(f)/S_j(f)$是作为一个有 $2N_B$ 自由度的χ随机变量的近似分布．用此结论对基于$\hat{S}_j^{(\mathrm{sa})}(\cdot)$的 A_j，导出一个近似无偏估计．随着 N_B 的变大，这个估计量将采用什么形式呢？ ◁

8.5 小波方差的谱估计

现在让我们考虑小波方差的估计如何可以转化成谱密度函数的估计，假设$\{X_t\}$是一个具有谱密度函数 $S_X(\cdot)$和样本区间 Δt 的过程，它隐含了 $S_X(\cdot)$是定义在区间$[-f_N, f_N]$上，其中 $f_N\equiv1/(2\Delta t)$是奈奎斯特频率．令 $\nu_X^2(\tau_j)$是物理尺度 $\tau_j\Delta t\equiv2^{j-1}\Delta t$ 上的小波方差．如 8.1 节所述，极大重叠离散小波变换的小波滤波器$\{\tilde{h}_{j,l}\}$的带通特性推出

$$\nu_X^2(\tau_j)\approx2\int_{1/(2^{j+1}\Delta t)}^{1/(2^j\Delta t)}S_X(f)\mathrm{d}f \tag{315}$$

(这是式(297a)考虑到不是 1 的样本区间的重新表示)．随着用于形成$\{\tilde{h}_{j,l}\}$的小波滤波器$\{\tilde{h}_l\}$的宽的增加，上面的近似改善了，因为$\{\tilde{h}_{j,l}\}$变成了一个理想的带

通滤波器的更好的逼近(最引人注目的是标称通带外的频率取消，参见图 122).

在假定选择 L 使式(315)成为一个合理的近似之下，我们可以用函数 $\bar{S}_X(\cdot)$ 估计 $S_X(\cdot)$，$\bar{S}_X(\cdot)$ 是每个区间 $\left[\frac{1}{2^{j+1}\Delta t}, \frac{1}{2^j\Delta t}\right]$($j=1, \cdots, J_0$)上的分段常数，即我们假设

$$\bar{S}_X(f)=C_j,\quad \frac{1}{2^{j+1}\Delta t}<f\leqslant\frac{1}{2^j\Delta t},$$

其中 C_j 是一个常数，定义使得

$$\int_{1/(2^{j+1}\Delta t)}^{1/(2^j\Delta t)} S_X(f)\mathrm{d}f=\int_{1/(2^{j+1}\Delta t)}^{1/(2^j\Delta t)} \bar{S}_X(f)\mathrm{d}f=\frac{C_j}{2^{j+1}\Delta t}.$$

利用式(315)，我们有

$$\nu_X^2(\tau_j)\approx\frac{C_j}{2^j\Delta t},$$

因此我们可以使用

$$\hat{C}_j\equiv 2^j\,\hat{\nu}_X^2(\tau_j)\Delta t \tag{316}$$

来估计谱密度函数的阶. 基于小波谱密度函数估计的两个例子将在图 330 中讨论.

给定 $\hat{\nu}_X^2(\tau_j)$，$j=1, \cdots, J_0$，以上方案给出了区间$[1/(2^{J_0+1}\Delta t), f_N]$上谱密度函数的一个估计，但是不适合于低频 $f\in[0, 1/(2^{J_0+1}\Delta t)]$. 因为 J_0 层极大重叠离散小波变换尺度系数与这个频率区间有关，如果假设$\{X_t\}$是一个平稳过程(参见式(302)和临近的讨论)，我们可以在估计 $\mathrm{var}\{\bar{V}_{J_0,t}\}$的基础上填好我们的估计 $S_X(\cdot)$. 如果我们知道 $E\{X_t\}=0$，继而还有 $E\{\bar{V}_{J_0,t}\}=0$，那么一个适当的估计是

$$\widehat{\mathrm{var}}\{\bar{V}_{J_0,t}\}\equiv\frac{1}{M_{J_0}}\sum_{t=L_{J_0}-1}^{N-1}\widetilde{V}_{J_0,t}^2;$$

然而，如果 $E\{X_t\}$是未知的，那么 $E\{\bar{V}_{J_0,t}\}$也是未知的(在实际中，通常也是这样的). 那么我们必须用

$$\widehat{\mathrm{var}}\{\bar{V}_{J_0,t}\}\equiv\frac{1}{M_{J_0}}\sum_{t=L_{J_0}-1}^{N-1}(\widetilde{V}_{J_0,t}-\overline{\widetilde{V}}_{J_0})^2,$$

其中$\overline{\widetilde{V}}_{J_0}$是观察的 $\widetilde{V}_{J_0,t}$的样本均值. 如果$\{X_t\}$是一个长记忆过程，那么它就是$\{\bar{V}_{J_0,t}\}$，在这种情况下我们必须意识到，因为上述估计是式(299a)的形式，它会严重低估 $\mathrm{var}\{\bar{V}_{J_0,t}\}$. 因此，如果过程均值是已知的或者如果$\{X_t\}$并没有展示长记忆特征，我们就应该可以通过采用$\widehat{\mathrm{var}}\{\bar{V}_{J_0,t}\}$的适当估计和用它形成谱密度函数 2^{J_0} 阶$\widehat{\mathrm{var}}\{\bar{V}_{J_0,t}\}\Delta t$ 去覆盖区间$[0, 1/(2^{J_0+1}\Delta t)]$来获得一个 $S_X(\cdot)$的合理估计.

8.5 节的评论与扩展

[1]如 8.1 节提到的，通过用哈尔小波连同式(315)得到 Blackman and Tukey (1958)的试验谱估计量(参见 7.3.2 节，Jenkins and Watts，1968)．在他们书的 18 节 Blackman 和 Tukey 描述了一个简单的试验估计量，然后在 B.18 节提出一些更改．在一个比例常数内，他们的谱水平 C_j 的简单估计量通过 $2^j \hat{\nu}_X^2(\tau_j)\Delta t$ 给出，其中$\hat{\nu}_X^2(\tau_j)$是式(308a)给出的基于离散小波变换的无偏估计量，除非特殊化为哈尔小波：

$$\hat{\nu}_X^2(\tau_j) = \frac{1}{N}\sum_{t=0}^{N_j-1} W_{j,t}^2,$$

而 $W_{j,t}=(V_{j-1,2t+1}-V_{j-1,2t})/\sqrt{2}$，$V_{j,t}=(V_{j-1,2t+1}+V_{j-1,2t})/\sqrt{2}$和 $V_{0,t}\equiv X_t$(把上面的与(308a)比较，对于哈尔小波 $L'_j=0$)．就像数值例子指出的那样(Blackman and Tukey，1958，p.46，表 III)，对于哈尔小波，他们实际上用离散小波变换塔式算法计算简单的试验估计(又一次忽略了比例常数)．

另外，简单试验估计的一个更改是一个“完全变型”(Blackman and Tukey，1958，p.136，表Ⅵ)，它与 $\tilde{\nu}_X^2(\tau_j)$的无偏哈尔极大重叠离散小波变换估计值成比例，并且对于哈尔小波通过一个接近极大重叠离散小波变换塔式算法的算法完成(算法差异仅仅有一个交叉运算和一个比例常数)．他们用一个紧密类似于 5.3 节用过的论证得到这个估计量，消除一个运算包括“丢掉每一个间隔值”(下抽样)．

最后，在下面的 8.6 节，我们探索哈尔小波方差 $\nu_X^2(\tau_j)$相反的会由线性趋势影响，在相关趋势情形下，我们建议用 $L\geqslant 4$ 的基于 Daubechies 小波滤波器的小波方差．Blackman and Tukey，(1958，p. 136)也认为试验估计量对于线性趋势很敏感，作为补救，建议差分化时间序列，然后把这些差分的数据进行实验分析，这个补救类似于在其里面用 D(4)小波与哈尔小波对比，D(4)滤波器有一个额外的嵌入差分运算．

8.6 例子：原子钟偏差

作为小波方差应用的第一个例子，让我们回到第 1 章所研究的一些数据，并考虑图 318 上面部分的时间序列，该图显示了时间$\{X_t\}$与铯原子钟(编号为“时钟 517”)及美国首都华盛顿美国海军气象台所保存的合作世界时间(美国海军气象台)所表现的官方时间尺度的差异．这种时间尺度作为一种“完美的”钟从而常用来研究单个时钟 571 的不完善性(事实上，合作世界时间(美国海军气象台)自身是由原子钟的全体平均形成的，当然不完善，它也被认为至少在保持时间上优于时钟 571 一个数量级)．同时在 20 世纪 70 年代连续 $N=1\ 026$ 天记录了每天时钟

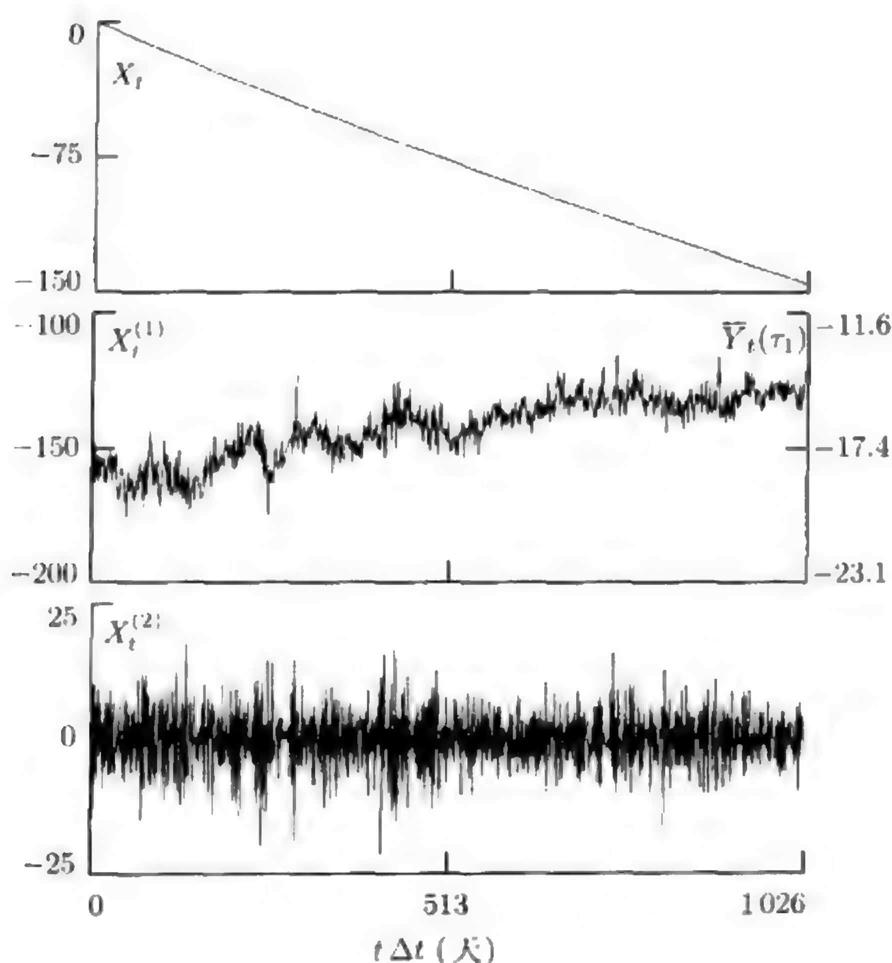


图 318 由时钟 571(一个铯原子钟)得到的和以美国首都华盛顿海军气象台保存的合作世界时间(美国海军气象台)为时间尺度得到的时间序列$\{X_t\}$的差异图(上部图)；它的一阶向后差分$\{X_t^{(1)}\}$(中间)；二阶向后差分$\{X_t^{(2)}\}$(下部). 在中间图中，$\bar{Y}_t(\tau_1)$表示τ_1平均分形频率偏离(由10^{13}的部分给出)——它们是由式(321c)定义的，与$X_t^{(1)}$成正比

571 和合作世界时间(美国海军气象台)在时间上的区别，因此样本的时间间隔$\Delta t=$ 1 天. 每个X_t的值是以毫微秒(1 秒$=10^9$毫微秒)为单位记录的，但为方便起见平面图的纵坐标以微秒(1 秒$=10^6$微秒)为单位表示. 如果在特殊的一天$X_t=0$，那就表示时钟 571 被设置的和合作世界时间(美国海军气象台)的时间保持完美的一致，这发生在$t=0$的那一天，仅仅是因为时钟 571 被设置的和合作世界时间(美国海军气象台)一致. 如果在第t天时钟 571 比合作世界时间(美国海军气象台)慢(例如，有一天合作世界时间(美国海军气象台)报时现在是中午 12 点以后，时钟 571 才显示中午 12 点)，那么习惯上称$X_t<0$. 图 318 给出了实际中的所有准确的值，因此在近三年的时期内时钟 571 会越来越落后合作世界时间(美国海军气象台)，最后大约落后 146 微秒.

为对时间序列估计小波方差，可以使用式(306b)的估计量$\hat{\nu}_X^2(\tau_j)$，首先我们需要考虑如何很好地把$\{X_t\}$各种向后差分看作均值为零的一个平稳过程的实现.

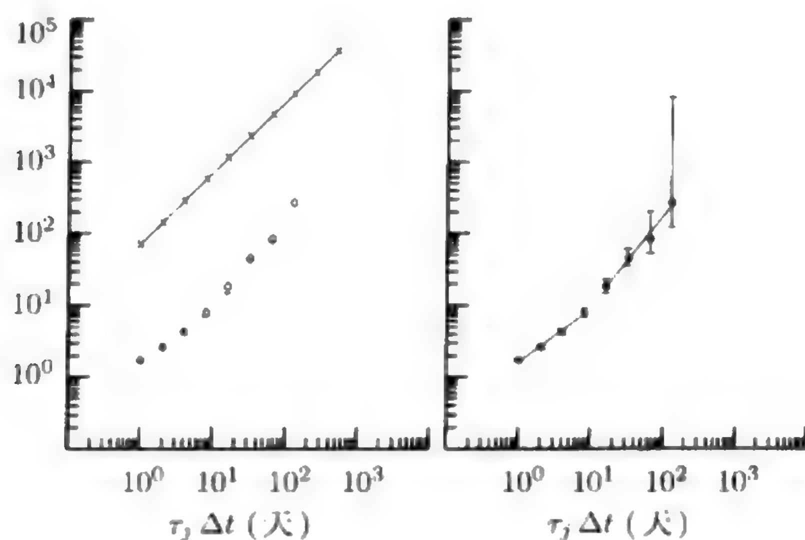


图 319　对于原子钟时间差分$\{X_t\}$，基于无偏极大重叠离散小波变换估计量和以下小波滤波器：哈尔(左边图中的 x，有一条最小二乘直线穿过它们)，D(4)(左右图中的圆圈)和 D(6)(左边图中的加号)的小波方差估计的平方根．第二个图中 95%的置信区间是通过当 $j=1,\cdots,6$ 时用由式(313d)给出的$\hat{\eta}_1$ 代替 η，当$j=7$，8时，用由式(314c)给出的 η_3代替 η，利用式(313c)计算的区间的平方根．不同的$\hat{\nu}_X(\tau_j)$的真实值在本书的网页中列出

图 318 的中图和下图显示了$\{X_t\}$的一阶和二阶向后差分，即

$$X_t^{(1)} \equiv X_t - X_{t-1}, X_t^{(2)} \equiv X_t^{(1)} - X_{t-1}^{(1)} = X_t - 2X_{t-1} + X_{t-2}$$

(注意现在纵坐标的单位是毫微秒)．第一个差分序列$\{X_t^{(1)}\}$的样本均值(−142.3)明显偏离 0，并且随时间有线性增大的趋势，因此认为这个序列是一个均值为零的平稳过程的实现是不合理的．第二个差分序列$\{X_t^{(2)}\}$和它的变化域相比有一个很小的样本均值(0.023)，而且没有明显的不连续，是一个平稳过程的实现；然而，如果$\{X_t^{(1)}\}$确实有线性趋势而有斜率 a_1，练习[288b]表示$\{X_t^{(2)}\}$的均值会等于 a_1．由于在一千天里$\{X_t^{(1)}\}$增加大约 25 毫微秒，那么 a_1 的粗略估计是 0.025，和样本均值相当一致．由于 $d=2$ 是减小$\{X_t\}$到小波方差有意义的最小差额的量，那么我们应该考虑用宽度为 $L\geqslant 2d=4$ 的小波滤波器．除非$\{X_t^{(1)}\}$的线性趋势的确有问题，否则 D(4)小波应该是适当的，在线性趋势有问题情况下我们需要用 D(6)小波(这和练习[305]的结果一致)．

图 319 的左图显示了对于 D(4)和 D(6)小波滤波器的估计的作为 τ_j 的函数的小波标准偏离$\hat{\nu}_X(\tau_j)$(即$\hat{\nu}_X^2(\tau_j)$的平方根)，说明一下，哈尔小波滤波器也是如此(在原子钟著作里画标准偏离图而不是方差是常见的，因为前者在研究中和数据的单位相同)．在左图中，对于 D(4)和 D(6)，$\hat{\nu}_X(\tau_j)$的值分别用圆圈和加号表示

出来；而对于哈尔小波滤波器用通过×的直线示出，即作为 $\log_{10}(\tau_j)$ 的函数的点 $\log_{10}(\hat{\nu}_X(\tau_j))$ 的最小二乘拟合，$j=1, 2, \cdots, 10$. 注意到估计的 D(4) 和 D(6) 小波方差相互之间非常接近，而且可以看出两个 $\hat{\nu}_X^2(\tau_j)$ 估计值相同(即 $S_X(\cdot)$ 在 $\frac{1}{2^{j+1}} < f \leqslant \frac{1}{2^j}$ 时的二次积分)，$\{X_t^{(1)}\}$ 的明显线性趋势对 D(4) 在 τ_7 处的估计值影响不大而且偏低.

练习[320a] 解释为什么一共有十个哈尔 $\hat{\nu}_X^2(\tau_j)$ 估计值，而分别仅有 8 个和 7 个估计值，估计 D(4) 和 D(6) 小波方差. ◁

虽然 D(4) 和 D(6) 小波方差相当一致，而哈尔小波方差截然不同，毫无疑问，因为 $\{X_t\}$ 的一阶向后差分被认为是一个零均值平稳过程的实现是不合理的. 通过哈尔 $\hat{\nu}_X(\tau_j)$ 值的最小二乘直线能很好拟合，斜率十分接近于 1(0.998)，这一点可用下面的练习解释.

练习[320b] 令 $\{X_t'\}$ 是一阶向后差分均值为零的平稳的过程. 定义 $X_t \equiv X_t' + a_0 + a_1 t$；即 $\{X_t\}$ 和 $\{X_t'\}$ 一样，但有附加的线性趋势. 证明，如果用哈尔小波滤波器，那么

$$E\{\hat{\nu}_X^2(\tau_j)\} = \nu_{X'}^2(\tau_j) + C\tau_j^2,$$

其中常数 C 取决于斜率 a_1，确定它是练习的一部分. 使用图 318 的上图可以粗略地得到 a_1，因此就可以确定 C，并且解释在图 319 中左图的哈尔 $\hat{\nu}_X(\tau_j)$ 值. ◁

注意甚至没有仔细考虑 $\{X_t\}$ 的特征，这容易直接从图 319 中看出，有时的现象用哈尔小波方差会出错，因为它和 D(4) 及 D(6) 方差很不一致. 一般地，用不同宽度的滤波器产生小波方差防止选出的 L 过小是一个不错的办法.

由于图 319 指出 D(4) 和 D(6) 小波方差的一点不同，我们可以安排 D(4) 估计足以确定时钟 571 的性能(然而，我们必须事先推断 D(4) 在 τ_8 处的估计，因为没有相应的 D(6) 估计，就像我们后面将看到的，由于 $\{X_t^{(1)}\}$ 的明显的线性漂移，D(4) 估计可能会存在偏差). 图 319 中的右图显示了 D(4) 小波方差估计的平方根与使用由式(313d)的 τ_1 到 τ_6 的 $\hat{\eta}_1$ 和式(314c)的 τ_7 和 τ_8 的 η_3 估计的自由度为 η 的 χ^2 逼近构造的小波方差的 95% 置信区间的平方根(后两个尺度有等价自由度 $\eta_3=3$ 和 1 相当于 $\hat{\eta}_1=13$ 和 3，无疑将产生乐观的置信区间). 例如，这个图告诉我们，$\tau_1=1$ 天时钟 571 的时间偏差的平均值的差接近于 2 毫微秒，在 95% 置信水平上，$\nu_X(\tau_1)$ 大至 3 毫微秒小至 1 毫微秒是不合理的. 这个分析可以用来比较时钟 571 和其他时钟的性能. 由于我们确定 $\hat{\nu}_X(\tau_1)$ 怎样能接近真值 $\nu_X(\tau_1)$，就可以以此来判断时钟 571 和其他任何时钟在性能上的观察区别是具有统计意义的，

或者仅仅是由于数据数量的限制的人为统计结果.

从图 319 也可以看出，随着 τ_j 的增大时钟 571 的平均时间差也在变化. $\hat{\nu}_X(\tau_j)$随 τ_j 变化的状态可看成是 $\log_{10}(\hat{\nu}_X(\tau_j))$作为 $\log_{10}(\tau_j)$的函数的两个区域线性增加的结果，即 τ_1 到 τ_4 和 τ_5 到 τ_8. 图 319 显示在这两个区域的线性最小二乘法拟合，τ_1 到 τ_4 的斜率为 0.74，τ_5 到 τ_8 的斜率为 1.26(τ_4 似乎是转折点，因为推断的第二个衰退线和$\hat{\nu}_X(\tau_4)$非常接近). 上述解释同有限频率的两个幂规律的情况一致，一个在小尺度(8 天或更小)，一个在大尺度(16 天或更大). 借助于$\nu_X(\tau_j)$，可以概括出时钟 571 的性能如下：

$$\hat{\nu}_X(\tau_j) \approx \begin{cases} 1.624\tau_j^{0.74}, & j = 1,2,3,4; \\ 0.541\tau_j^{1.26}, & j = 5,6,7,8. \end{cases} \tag{321a}$$

注意借助于 $\nu_X^2(\tau_j)$而不是它的平方根，上式变为 $\nu_X^2(\tau_j) \propto \tau_j^{1.48}$，$j=1$，2，3，4 和 $\nu_X^2(\tau_j) \propto \tau_j^{2.52}$，$j=5$，6，7，8. 因为 $\nu_X^2(\tau_j) \propto \tau_j^{-\alpha-1}$ 转移为 $S_X(f) \propto |f|^{\alpha}$，$\frac{1}{2^{j+1}} < |f| \leqslant \frac{1}{2^j}$，所以，近似地有

$$S_X(f) \propto \begin{cases} |f|^{-2.48}, & 1/32 < |f| \leqslant 1/2; \\ |f|^{-3.52}, & 1/512 < |f| \leqslant 1/32. \end{cases} \tag{321b}$$

现在让我们来讨论哈尔小波方差和 Allan 方差之间的关系. Allan 方差是 Allan(1996)提出的一个不错的时钟性能测量(参见 Barnes 等，Rutman，1978；Percival，1983，1991；Greenhall，1991；Flandrin，1992；以及 Percival and Guttorp，1994). 为了建立这种关系，我们从对 τ_1 的平均分形频率偏离的定义开始

$$\bar{Y}_t(\tau_1) \equiv \frac{X_t - X_{t-1}}{\tau_1 \Delta t}, \tag{321c}$$

它和$\{X_t\}$的一阶向后差分成比例. 如果$\{X_t\}$的单位用毫微秒，$\tau_1\Delta t$ 的单位用天，那么如果 $\tau_1\Delta t$ 再用毫微秒为单位，$\bar{Y}_t(\tau_1)$可变为无量纲的量. 因为一天是 86 400 秒，在这个例子中可以给一阶差分 $X_t - X_{t-1}$除以 8.64×10^{13}使 $\bar{Y}_t(\tau_1)$变为无单位的. 一天里 X_t 中的 8.64 毫微秒的改变相当于$\frac{1}{10^3}$的分形频率偏离(参见图 318 中间图的纵轴).

因为一个时钟仅仅是一个按照振荡工作在一个特殊额定频率 f_0(例如，$f_0=$ 5 兆赫)下循环计数的装置，因此平均分形频率偏离的概念是有用的，因为它告诉了在一个时间周期 $\tau_1\Delta t$ 里，在相对的基础上额定频率偏离振荡器多少. 特别地，如果令 $f(t)$在 t 时刻是振荡器的“瞬时”频率(这和相位变化的时间导数有

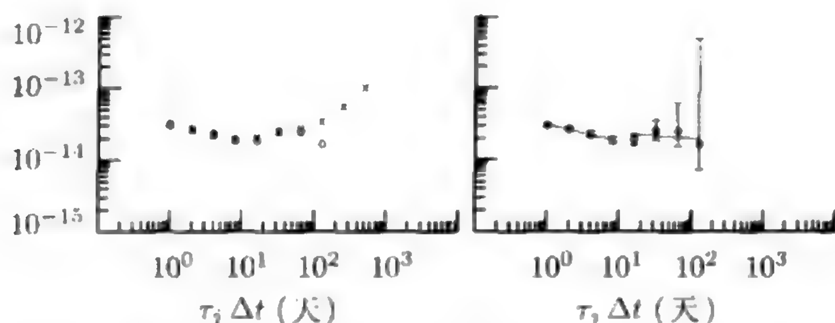


图 322 对于基于无偏极大重叠离散小波变换估计量和以下小波滤波器：哈尔（左边图中的×）和 D(4)（左右两边图中的圆圈）的原子钟一天平均分形频率偏离$\{\bar{Y}_t(\tau_1)\}$的小波方差估计的平方根．$\hat{\nu}_Y(\tau_j)$的真实值在本书的网页中列出

关)，那么我们有

$$\bar{Y}_t(\tau_1)=\frac{1}{\tau_1\Delta t}\int_{t-\tau_1\Delta t}^{t}\frac{f(s)-f_0}{f_0}\mathrm{d}s.$$

同样，τ_j 平均分形频率偏离可以给出如下：

$$\bar{Y}_t(\tau_j)\equiv\frac{1}{\tau_j\Delta t}\int_{t-\tau_j\Delta t}^{t}\frac{f(s)-f_0}{f_0}\mathrm{d}s=\frac{1}{\tau_j}\sum_{l=0}^{\tau_j-1}\bar{Y}_{t-l}(\tau_1)=\frac{X_t-X_{t-\tau_j}}{\tau_j\Delta t}.$$

特别注意，$\bar{Y}_t(\tau_j)$是 τ_1 偏离子序列的平均，因此是$\{\bar{Y}_t(\tau_1)\}$的函数.

如果现在我们把$\{\bar{Y}_t(\tau_1)\}$看作是一个随机过程，它的一阶向后差分是一个平稳过程，那么在尺度 τ_j 处对于这些 τ_1 偏离，Allan 方差定义为

$$\sigma_Y^2(2,\tau_j)=\frac{1}{2}E\{(\bar{Y}_t(\tau_j)-\bar{Y}_{t-\tau_j}(\tau_j))^2\}\tag{322}$$

(练习[8.12]在 Allan 方差注释后给出了基本原理，而练习[8.13]解释了为什么在期望前有一个因子$\frac{1}{2}$)．现在可以把$\{\bar{Y}_t(\tau_1)\}$的 Allan 方差与哈尔小波方差联系如下.

练习[322] 令$\{\bar{Y}_t(\tau_1)\}$是具有平稳一阶向后差分的过程．令 $\nu_{\bar{Y}}^2(\tau_j)$是这个过程的哈尔小波方差．证明，如果向后差分过程均值为零，那么

$$\nu_{\bar{Y}}^2(\tau_j)=\frac{1}{2}\sigma_Y^2(2,\tau_j).$$ ◁

让我们考虑，以对$\{X_t\}$前面的分析和重新改变比例一阶向后差分$\{\bar{Y}_t(\tau_1)\}$有什么关系．图 322 显示了对于$\{\bar{Y}_t(\tau_1)\}$估计哈尔和 D(4)小波方差的平方根(分别在图中由×和圆圈给出)．因为现在处理的是过程$\{\bar{Y}_t(\tau_1)\}$，它的一阶差分和均值为零

的平稳过程接近哈尔和 D(4)小波方差现在是完全可以比较的，唯一明显的差别是在尺度 $\tau_8=128$ 天，其中 D(4)的估计约比哈尔小波方差小三倍．如果用哈尔估计计算出最小二乘法拟合与式(321a)比较，可以得到

$$\hat{\nu}_Y(\tau_j)\approx\begin{cases}3.093\times10^{-14}\tau_j^{-0.23}, & j=1,2,3,4;\\ 1.031\times10^{-14}\tau_j^{0.24}, & j=5,6,7,8.\end{cases}\tag{323}$$

转换后得

$$S_Y(f)\propto\begin{cases}|f|^{-0.54}, & 1/32<|f|\leqslant 1/2;\\ |f|^{-1.48}, & 1/512<|f|\leqslant 1/32.\end{cases}$$

因为 $\{\overline{Y}_t(\tau_1)\}$ 是和 $\{X_t\}$ 的一阶差分成比例，我们有

$$S_X(f)\propto\frac{S_Y(f)}{\sin^2(\pi f)}\approx\frac{S_Y(f)}{(\pi f)^2}\propto\begin{cases}|f|^{-2.54}, & 1/32<|f|\leqslant 1/2;\\ |f|^{-3.48}, & 1/512<|f|\leqslant 1/32.\end{cases}$$

上面的指数和式(321b)一致．

此外，如果对于两个最大尺度我们取哈尔估计并确定通过它们的直线(在 log/log 空间)，对于 $j=9$ 和 10 求得 $\hat{\nu}_Y(\tau_j)\propto\tau_j^{0.88}$，因此指数相当接近于 1．练习[320b]讨论指数能够为 1 是由于在 $\{\overline{Y}_t(\tau_1)\}$ 里的主线性趋势所造成的，这在图 318 的中已经注意到了．线性趋势不应该引起对于 D(4)小波方差的问题，因为 D(4)滤波器有两个嵌入差分运算．由于哈尔估计 $\hat{\nu}_Y(\tau_8)$ 位于十分接近基于拟合尺度 τ_9，τ_{10} 外推，观察对于 $\nu_Y(\tau_8)$ 的哈尔和 D(4)估计的区别可能是由于线性趋势的缘故；即由于这个趋势哈尔估计可能上偏．图 322 中的右图显示了由式(313d)对尺度 τ_1 到 τ_6 得到的 $\hat{\eta}_1$ 和由式(314c)的尺度 τ_7 到 τ_8 得到的 η_3 估计是得到用式(313c)确定置信度为 95%的 D(4)的置信区间．哈尔估计 $\hat{\nu}_Y(\tau_8)$ 落在基于 D(4)估计的对于 $\nu_Y(\tau_8)$ 的置信区间内，因此样本变化可能也会造成差异．

最后，让我们用这个例子说明小波方差图可以帮助确定，当一个时间序列如同一个平稳过程一样对着一个有平稳向后差分的不平稳过程的最好的模型．按照 $\frac{1}{f}$ 型过程，前者通过在低频部分随着 $S_X(f)\propto|f|^\alpha$，$\alpha>-1$ 近似用谱密度函数变化特征化，而后者在 $\alpha\leqslant-1$ 特征化．对于小波方差，在大尺度下，我们有近似的 $\nu_X^2(\tau_j)\propto\tau_j^{-1-\alpha}$，因此平稳性与负指数有关，并且不平稳过程与非负指数有关．当我们把 $\log_{10}(\nu_X^2(\tau_j))$ 作为 $\log_{10}(\tau_j)$ 的函数绘图时，指数变成了斜率，且在大尺度下零斜率区分为平稳过程和非平稳过程变成一个合适的参考．因此分形频率偏离 $\{\overline{Y}_t(\tau_1)\}$ 与一个非平稳模型保持一致．因此，在图 319 中标出的正斜率立即告诉我们时间差分 $\{X_t\}$ 是保持非平稳模型．图 322 左边的哈尔估计显示了大尺度下的正斜率，因此分形频率偏离 $\{\overline{Y}_t(\tau_1)\}$ 保持非平稳模型；然而，这种模型的需要仅仅由于一个线性趋势——D(4)估计值(它被这样一个趋势影响)提出一个平稳

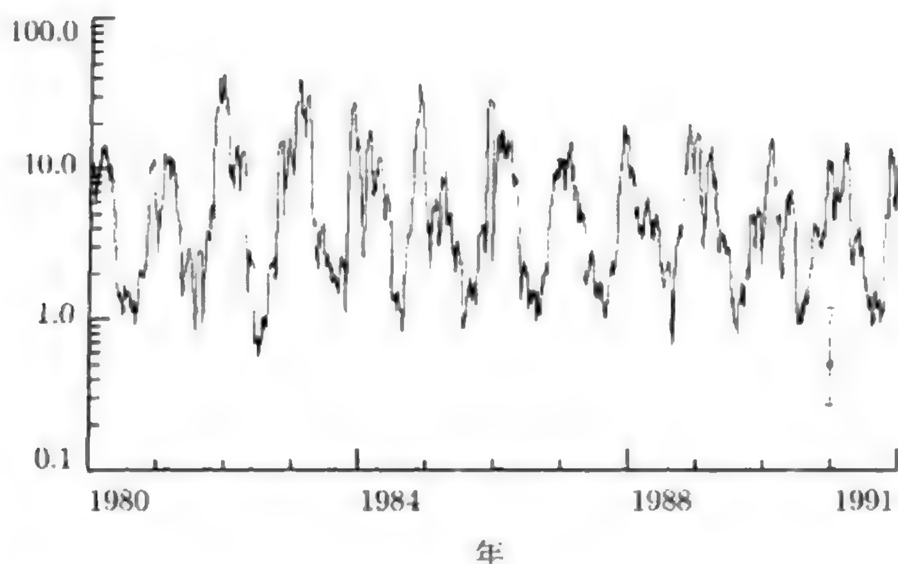


图 324 图 186 中新奥尔良市子潮海平面的变化，在实际尺度 $\tau_2\Delta t=1$ 天下作出的依赖于时间的 LA(8)小波方差估计，同时还有基于假想小波方差估计值为 1/2 和有着 $\nu=15.25$ 的 χ^2 分布的一个典型的 95%置信区间(详情参见正文)

模型也许已经很充分，就像图 332 右图中描述的通过线性最小二乘直线的负斜率指出的.

我们将进而在 9.7 节检查原子钟数据，作为一个可以作为长记忆过程模型的时间序列的例子.

8.7 例子：子潮海平面涨落

在 5.8 节我们看到新奥尔良市子潮海平面变化的多分辨分析指出在一些尺度下的变化是依赖季节的(参见图 186). 在假设围绕大约三十天有价值数据的时间段内之下，观察到的极大重叠离散小波变换小波系数可以看作是一个零均值高斯平稳过程的实现，通过计算与时间有关的小波方差估计值进而探究可变性(具有这些数据，由经验可知这些假设都是合理的).

图 324 表明基于有 61 个数据值的一段上，对于尺度 $\tau_2\Delta t=1$ 天的 LA(8)极大重叠离散小波变换小波方差估计(由于 $\Delta t=\frac{1}{2}$天，每一个时间段跨越 30.5 天). 我们可以直接计算，对于 $j=2$，$N_S=61$，

$$\hat{\nu}^2_{X,t}(\tau_j)\equiv\frac{1}{N_S}\sum_{u=-(N_S-1)/2}^{(N_S-1)/2}\widetilde{W}^2_{j,t+|\nu_j^{(H)}|+u \bmod N}, \tag{324}$$

然后以与 X_t 有关的实际时间为横坐标，绘出这个估计的图(在上面，$|\nu_2^{(H)}|=11$ 是为了排列$\{\widetilde{W}_{2,t}\}$，LA(8)滤波器的相位特性所需的循环移位，因此，我们可以认为它是一个零相位滤波器的近似输出——详情参见 4.8 节). 在我们的假设

下，可以为基于随机变量$\hat{\nu}^2_{X,t}(\tau_2)$分布的 χ^2 逼近，假定随时间变化的真实小波方差 $\nu^2_{X,t}(\tau_2)$计算一个 95%置信区间. 我们可以用一个类似于产生式(314c)中 η_3的论证为 χ^2 分布置等价自由度是 $\eta=15.25$：每一个估计量用 61 个极大重叠离散小波变换小波系数和 $j=2$ 形成. 因此，置信区间采用形式

$$\left[\frac{\eta\hat{\nu}^2_{X,t}(\tau_2)}{Q_\eta(0.975)},\frac{\eta\hat{\nu}^2_{X,t}(\tau_2)}{Q_\eta(0.025)}\right];$$

在对数尺度下绘图时，我们有效地对 $\log\nu^2_{X,t}(\tau_2)$获得一个置信区间，形式为

$$\left[\log(\hat{\nu}^2_{X,t}(\tau_2))+\log\left(\frac{\eta}{Q_\eta(0.975)}\right),\log(\hat{\nu}^2_{X,t}(\tau_2))+\log\left(\frac{\eta}{Q_\eta(0.025)}\right)\right].$$

这个区间的宽是 $\log(Q_\eta(0.975)/Q_\eta(0.025))$，它明显与真正的估计值无关. 图 324 的右下角附近描述了基于假设估计$\hat{\nu}^2_{X,t}(\tau_2)=1/2$上的置信区间. 因为这个图有一个 $\log_{10}$的纵坐标，我们仅仅需要整体移动这个置信区间，以使低置信限和高置信限之间的圆圈以真实的估计$\hat{\nu}^2_{X,t}(\tau_2)$为中心，为 $\nu^2_{X,t}(\tau_2)$获得一个 95%置信区间(绘在一个对数轴上).

对于随着时间变化的小波方差的置信区间表明，当我们说明抽样变化时，发生在同日不同年的$\hat{\nu}^2_{X,t}(\tau_2)$几乎没有明显的变化；即可以把图 324 中的曲线模拟成周期是一年的周期函数. 对于尺度 τ_3 到 τ_7，一个类似的简化可以得到证明. 伴随着图 186 的多分辨分析的研究，这些结果提供了一个合理方法，总结与尺度有关的方差每一年的变化，通过平均 12 年会在一个给定月份下降的循环平移的极大重叠离散小波变换小波系数的平方. 这个平均的结果在图 326 中展现出来. 在简化假设每一个月的小波系数可以认为是一个平稳过程实现的一部分且今年和下一年的系数彼此无关下，8.4 节描述的方法很容易地适用于为潜在的真实方差构造一个大约 95%的置信区间. 如果我们忽略出现的很小的差异，因为日历的月有不同的天数，在对数尺度下这些区间的宽度又一次依赖于特殊的尺度，但是，在一个给定的尺度下，不依赖于方差估计本身. 因此，基于 11 月估计，在图 326 中对每个尺度给出的置信区间可以整体转移到其他月份，通过一个给定的尺度去评价抽样变化.

我们能够从图 326 和图 186 得到以下结论：尽管在不同的尺度，对于给定的月份，方差估计是不相同的，但是方差的暂时模式对于所有 6 个尺度是非常类似的(对于尺度 $\tau_7\Delta t=32$ 天它是成立的这个事实，在图 186 的多分辨分析中并不是很明显). 平均方差一般在冬季最高(11 月到 3 月)，在 7 月和 8 月最低，大约是在一年的最低季节海平面后的 3 到 4 个月(参见图 186 中的$\widetilde{\mathcal{S}}_7$). 平均冬天的方差

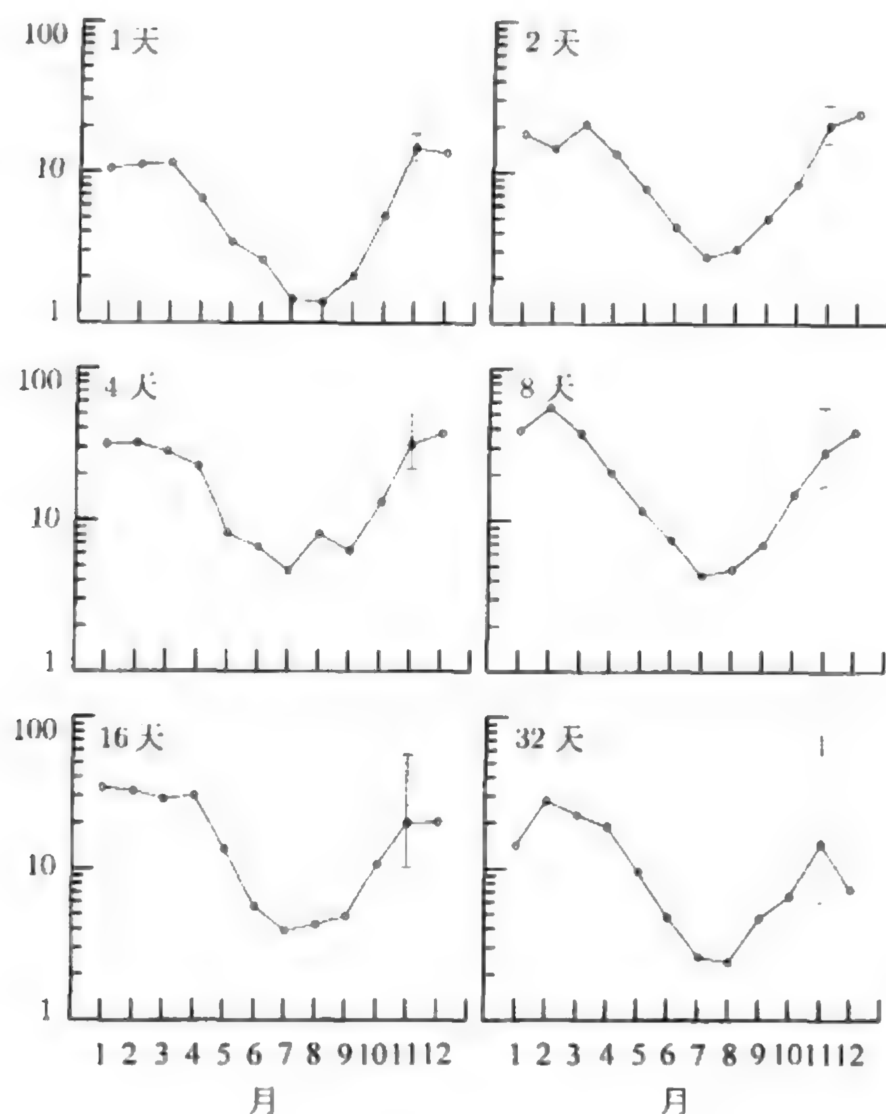


图 326 对于图 186 的子潮海平面的变化以日历的月份为组，在实际尺度 $\tau_j\Delta t = 2^{j-2}$ ($j = 2, \cdots, 7$) 天下估计 LA(8) 小波方差

比夏天的方差数量级高，这个相对简单的子潮海平面季节变化的基于尺度的描述能够使用，例如，可以用来通过计算机模拟研究它们的极值统计学——这个统计学在预测海啸影响的“最坏情况”中是非常重要的组成部分(Percival and Mofjeld, 1997).

8.8 例子：尼罗河最低水位

在 5.9 节尼罗河最低水位序列的讨论中，我们注意到大约在 715 年后，基于哈尔多分辨分析在视觉上提出在尺度 $\tau_1 = 1$ 和 $\tau_2 = 2$ 年下可变性降低，它是围绕着时间，历史记录显示在此时间内这个序列测量方式上的变化. 因为人类的视觉系统已经在某种程度上发展，因此，我们可以很容易地分辨出模式，当随机性提出时，会容易被欺骗从而看见不存在的模式. 因此，通过考虑抽样变异性的影响的统计过程，证实我们的视觉影响是很重要的. 为此，我们把时间序列分成两部分，由序列的前 94 个观测值(622 到 715)和后 569 个观测值(716 到 1 284)组成.

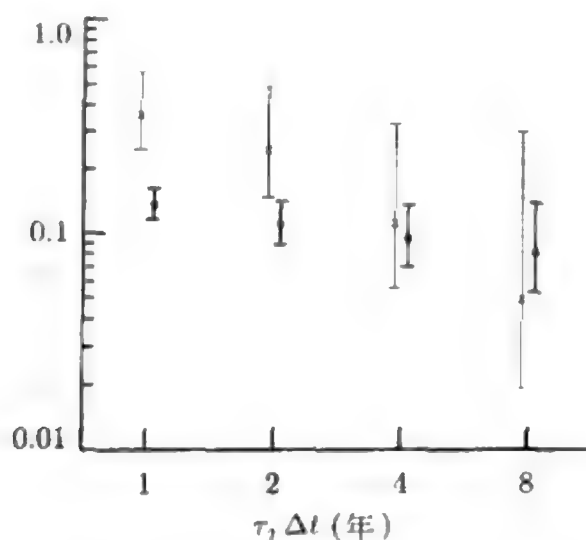


图 327　在 715.5 年前后对尼罗河最低水位时间序列估计的哈尔小波方差(分别用×和 o 表示)，还有基于有着由式(314c)中的 η_3 确定的等价自由度的 χ^2 逼近的 95%置信区间(分别用细线和粗线表示). $\hat{\nu}_X^2(\tau_j)$的真实值在本书的网站中列出

有了这个分割，然后我们在尺度 τ_1 到 τ_4 下单独为这两个子序列计算哈尔小波方差. 这些估计的方差在图 327 中给出，一起的还有用由式(314c)中使用 η_3决定的等价自由度的使用 χ^2 逼近计算出的 95%置信区间(在 8.4 节讨论的结果中，这个方法产生最宽的置信区间，因此被采用，以致于我们在留有余地方面犯了错误). 我们看到，对尺度 τ_1 和 τ_2，95%的置信区间确实没有捕获相同的值，导致对哈尔多分辨分析的视觉解释产生信任. 相反，对于尺度 τ_3 和 τ_4，95%置信区间几乎没有显示小波方差在这些尺度下有明显的改变.

这个例子说明了在一个并不是全局平稳，但可以认为在两个邻接的时间段内是局部平稳的序列上小波方差的用途，历史上的记录允许我们在 715 年处客观地把这个序列分开，但是在一个序列中区域的先验的辨认平稳性是合理的假设常常是不可能的. 就像在这里做的，在 9.6 节我们为不需要一个先验的分割的方差齐性做尺度统计测试(当应用于在 9.8 节的尼罗河序列时，我们得到与这里相同的结论).

8.9　例子：海洋切变测量

这里，我们通过把小波方差应用于 5.10 节描述的和图 194 下方绘出的垂直海岸线深度序列来说明它的用途. 在图上，有两条细的垂直线作标记，它们之间有从 489.5m 到 899.0m 的 4 096 个值. 随后，我们假设这个子序列(在图 328 的顶部再绘出)可以认为是一个过程实现的部分，这个过程的一次向后差分是一个

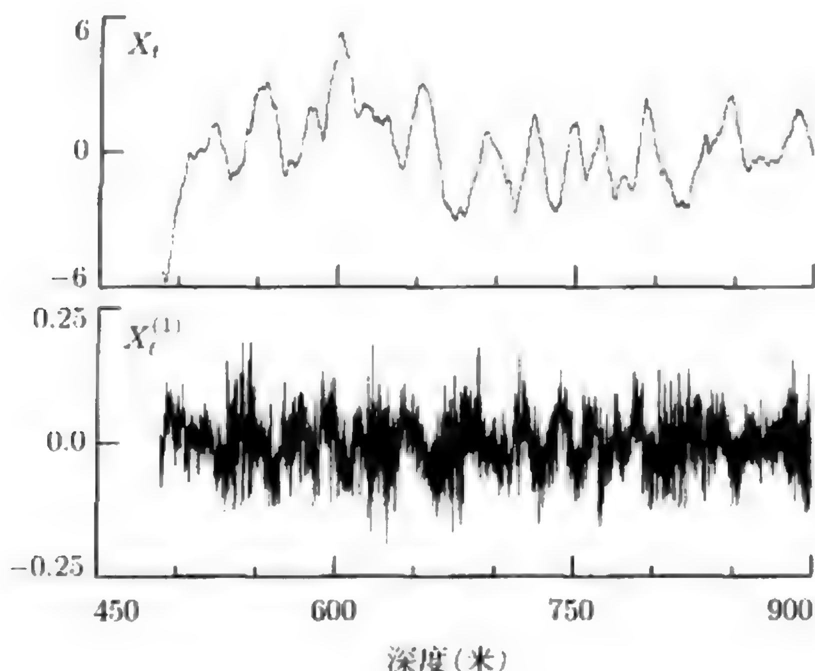


图 328 选择 $N=4\ 096$ 的海洋切变测量的部分$\{X_t\}$(上部的图)和相应的向后差分$\{X_t^{(1)}\}$(下部). 整个序列在图 194 的底部绘出，在它上面子序列由两条细的垂直线描绘

零均值的平稳过程(我们需要这个假设去应用下面用到的有意义的方法). 图 328 下方的图显示了这些差异，在图中没有任何显示它会排除一个零均值平稳过程作为模型.

图 329 左边图的×显示了在从 $\tau_1\Delta t=0.1$m 上升至 $\tau_{12}\Delta t=204.8$m 的实际尺度上的无偏极大重叠离散小波变换哈尔小波方差估计. 理论上，在一个 log/log 尺度下，作为 τ_j 函数的 $\nu_X^2(\tau_j)$图中的线性区域指出，在一个特殊频率区域上，有着与直线斜率相联系的幂规律指数的一个幂规律过程的存在. 对于最小的七个尺度，哈尔小波方差估计近乎完美地沿着这条直线下降. 图中穿过它们的直线是通过线性最小二乘计算的，其斜率为 $1.66\doteq5/3$. 由于对于一个有着指数 α 的幂规律过程，$\nu_X^2(\tau_j)$近似地随着 $\tau_j^{-\alpha-1}$ 变化(参见 8.1 节)，哈尔小波方差图有力地提出了在尺度 0.1 到 6.4 米上，有一个指数 $\alpha\doteq-8/3$ 的幂规律过程的存在.

除了哈尔估计外，图 329 左边的图也显示了基于 D(6)小波滤波器上的小波方差(+). 在小的尺度下，D(6)估计有规则地比哈尔估计小. 对于最小的七个尺度，D(6)估计值不再像一个明显的直线那样排成一行，并且对于相应的 D(6)估计，我们发现用哈尔小波方差最小的七个尺度的斜率看起来很不合理. 右边的图复制了 D(6)估计，也显示了基于 D(4)和 LA(8)小波滤波器上的估计(各自用小的和大的圆圈). 这里我们看到在所有的尺度下，D(6)和 LA(8)估计之间的一个好的吻合，并且排除在中间尺度下一些小的差异，它们和 D(4)估计值相当地

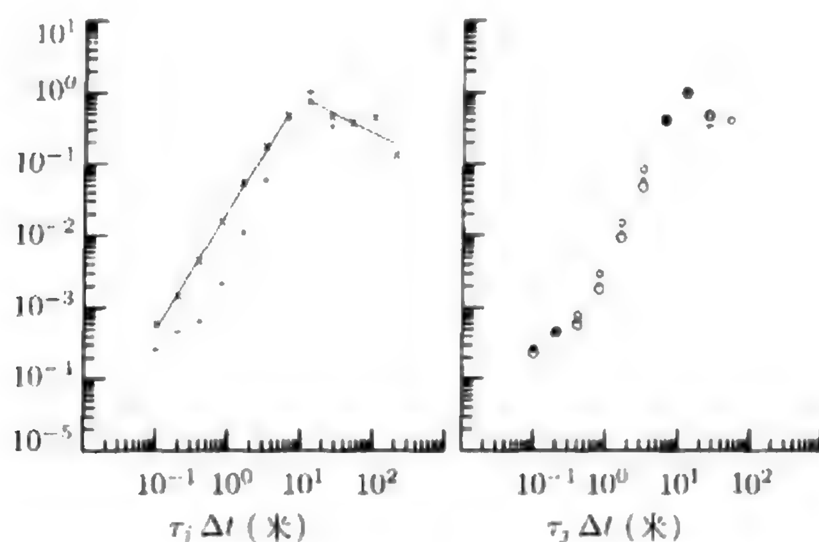


图 329 用无偏极大重叠离散小波变换估计量和下述小波滤波器得到的垂直切变序列的小波方差估计：哈尔(左边图中以×表示，有两条回归的直线拟和它们)，D(4)(右边图中的小圆圈)，D(6)(两个图中+所表示的)和 LA(8)(右边图中的大圆圈)．不同的$\hat{\nu}_X^2(\tau_j)$的值在本书的网页中列出

吻合．因此，虽然 D(4)、D(6)和 LA(8)估计合理地吻合，但是哈尔小波方差估计在小尺度下系统地比较高．在前节原子钟例子中，我们也注意到了，包含哈尔估计在内的不和谐性．我们认为这些不和谐性是由于对于原子钟数据向后差分不能合理地模拟成一个零均值平稳过程，从图 318 中间图可以明显地看到(样本均值实质上不等于零，并且可论证的是在差分的数据中有一个线性趋势)．对于垂直切变序列，哈尔和其他小波方差估计间的不一致不是由于任何明显地违背平稳性，但是可以解释为泄漏．

在谱密度函数估计中，泄漏是一个已经知道的现象——参见 7.5 节(特别是关于图 273 和 275 的讨论)．泄漏表明它自己在谱密度函数的低幂部分是一个向上的偏离．由于我们可以把小波方差估计转化为谱密度函数的一个区域估计，让我们这样做去说明提高的哈尔估计可以转化成与典型的泄漏形式符合的谱密度函数阶．图 330 显示了式(316)估计的谱密度函数阶 $\hat{C}_j$ 为一条在倍频程带 $\left[\frac{1}{2^{j+1}\Delta t}, \frac{1}{2^{j}\Delta t}\right]$上的常数直线，其中对于哈尔估计，$j=1$，…，12(粗“阶梯”)，对于 D(6)估计，$j=1$，…，9(细阶梯)．虽然在低频部分，两个谱密度函数估计之间达到一个好的吻合，但是在高频部分基于哈尔的估计偏高，它是估计的谱密度函数的低幂部分，这个模式与基于哈尔的估计中的泄漏相符合．

作为比较，我们也绘出了基于泄漏倾向周期图$\hat{S}^{(p)}(\cdot)$的谱密度函数估计(圆圈)和基于$K=7$正弦锥形和相对无泄漏的多锥估计量$\hat{S}^{(mt)}(\cdot)$(星号)(多锥谱密度函数估计量的详细情况参见 7.5 节)．在非零的傅里叶频率 $f_k \equiv \frac{k}{N\Delta t}$，

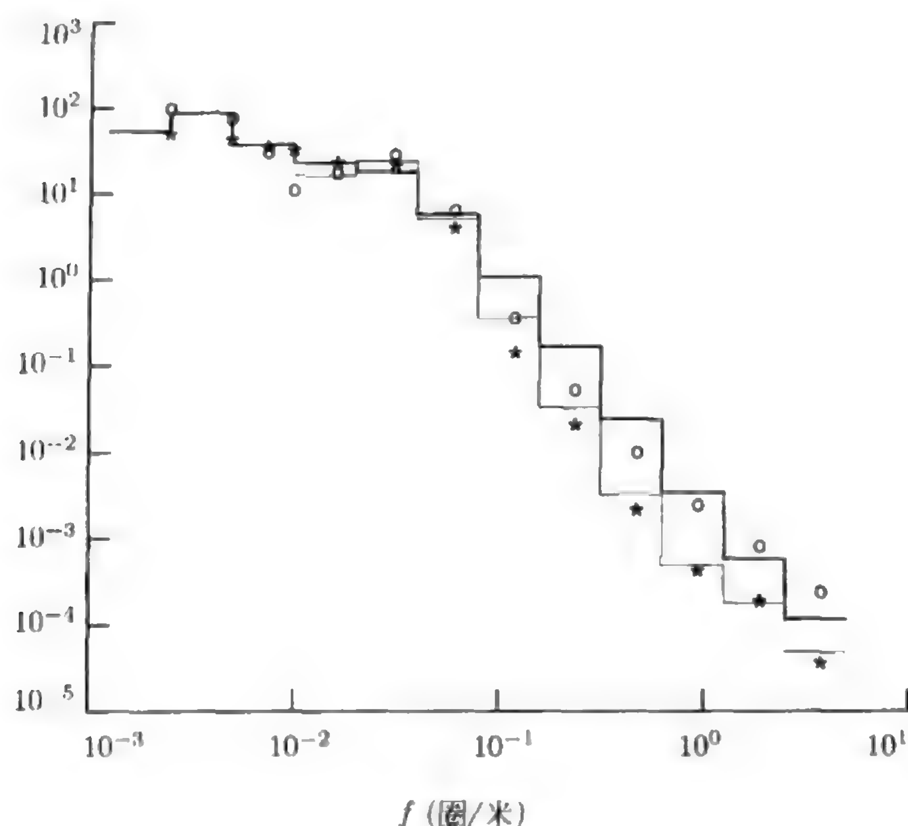


图 330　比较基于周期图(o)，用 $K=7$ 的正弦锥形形成的多锥谱密度函数估计(⋆)，哈尔和 D(6)小波方差估计(分别用粗的阶梯和细的阶梯表示)上的垂直切变测量"倍频程带"谱密度函数估计. 详情参见正文

$k=1, \cdots, 2\,048$(回想，$N=4\,096$)上计算周期图和多锥估计. 13 个圆圈表示，从左到右，$\hat{S}^{(p)}(f_1)$，$\hat{S}^{(p)}(f_2)$，$\hat{S}^{(p)}(f_3)$和$\hat{S}^{(p)}(f_4)$；下面四个估计值，即$\hat{S}^{(p)}(f_5)$，…，$\hat{S}^{(p)}(f_8)$的平均；下面八个估计值$\hat{S}^{(p)}(f_9)$，…，$\hat{S}^{(p)}(f_{16})$的平均；等等，以$\hat{S}^{(p)}(f_{1\,025})$，…，$\hat{S}^{(p)}(f_{2\,048})$的平均结束. 作为与估计值有关的傅里叶频率的函数，绘出每一个倍频程带平均的估计值的图. 星号是通过对多锥估计值按同样的方法获得. 注意在谱密度函数的低幂(高频)部分，基于周期图估计一贯地超过基于多锥估计，表明在周期图中有泄漏迹象. 通常，D(6)小波方差产生一个与摆脱泄漏的多锥估计值吻合很好的谱密度函数估计(在中间区域倍频程带中有一些差异，部分归因于抽样变化). 相反，基于哈尔谱密度函数估计显示了与周期图相同的模式；即在高频部分比 D(6)或基于多锥估计有规则地高. 再次，基于哈尔小波谱密度函数估计和泄漏倾向周期图性质上的类似表明在哈尔小波方差中一个泄漏形式.

为了确切地了解为什么基于哈尔谱密度函数估计容许泄漏，假定$\{X_t\}$有如下所述的理论谱密度函数，这粗略地类似于图 333 中的 D(6)和多锥估计：

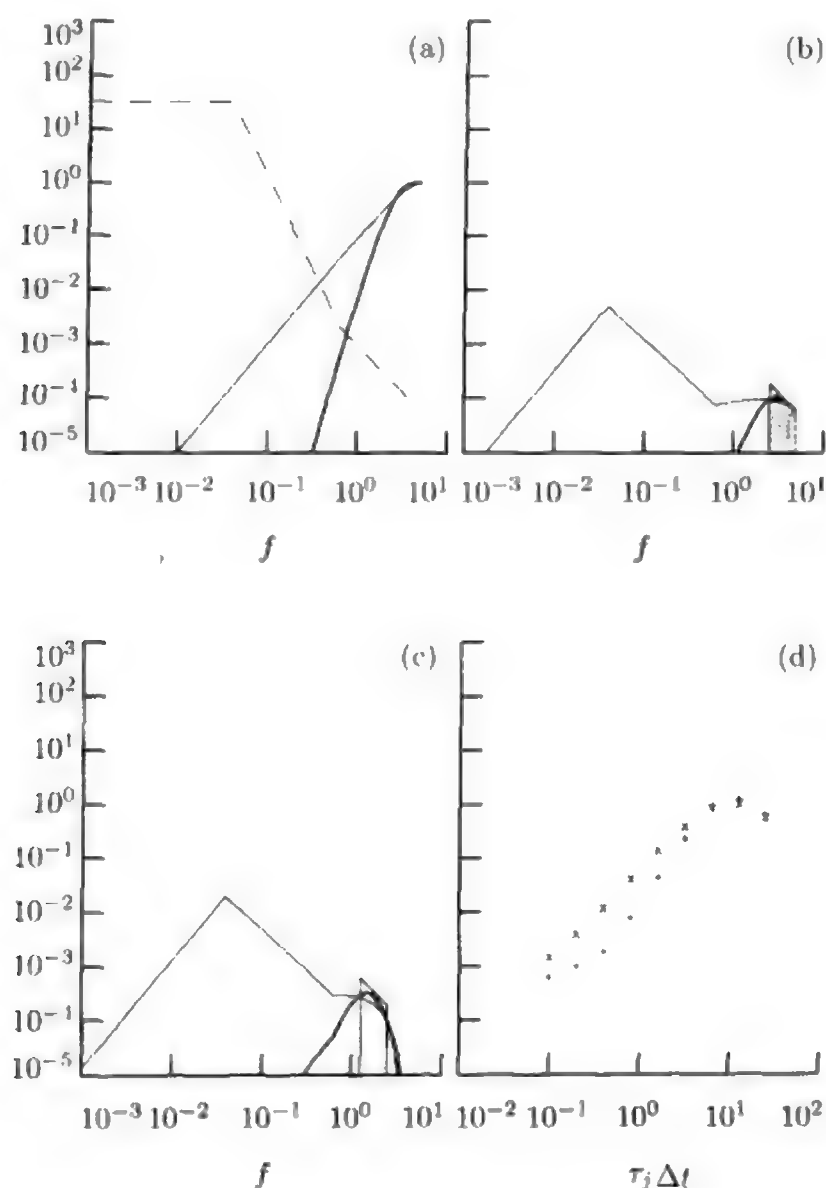


图 331　哈尔小波方差中的泄露(详情参见正文)

$$S_X(f)=\begin{cases}32.768, & |f|\leqslant 5/128;\\ 32.768\cdot|128f/5|^{-3.5}, & 5/128<|f|\leqslant 5/8;\\ |8f/5|^{-1.7}/500, & 5/8<|f|\leqslant 5.\end{cases}\tag{331}$$

图 331 中的图(a)表明 $S_X(\cdot)$(虚线)，一起的还有哈尔的平方增益函数 $\mathcal{H}_1^{(D)}(\cdot)$(细实线)和 D(6)(粗实线)小波滤波器. 图(b)的阴影部分描述了 $S_X(\cdot)$和对于一个理想的具有通带[5/2，5]的带通滤波器的平方增益函数的乘积，即标称上与物理尺度 $\tau_1\Delta t$ 有关. 这个图也显示了分别对于哈尔(细线)和 D(6)(粗)小波滤波器的 $S_X(\cdot)$和$\mathcal{H}_1^{(D)}(\cdot)$的乘积. 如果我们回想 $\nu_X^2(\tau_1)$是 $\widetilde{\mathcal{H}}_1^{(D)}(\cdot)$和 $S_X(\cdot)$乘积的积分(参见式(305))，我们看到在图(b)中细线和粗线下面的积分分别产生，对于最小的尺度的哈尔和 D(6)的小波方差. 虽然对于 D(6)小波的曲线很好地集中在标称通带附近，但是对于哈尔小波的曲线不是这样. 这个不同行为

的原因可由图(a)推断：对于小的 f，哈尔和D(6)小波的平方增益函数 $\widetilde{\mathcal{H}}_1^{(D)}(\cdot)$分别与 f^2 和 f^6 成比例(这些衰退率分别地与这些小波滤波器具有一个和3个嵌入的差分运算有关). 因此，哈尔衰退率不能补偿 $S_X(\cdot)$在频率的中间区域向上发散的快速率($f^{-3.5}$)，然而D(6)可以. 类似地，图(c)显示了第二个尺度的情形，它又一次显示了对于这个尺度，哈尔小波方差主要由标称通带外的频率决定. 最后，图(d)显示了前9个尺度的理论哈尔(×)和D(6)(+)小波方差(通过数值积分计算). 整体的现象和斜率与图329的左图符合得很好.

在图329中，在小尺度下，哈尔和其他小波方差的差异可归于泄漏的一种形式，泄漏是由

[1]对于这些尺度，标称通带外的哈尔平方增益函数的低衰退率，和

[2]对于垂直切变序列的谱密度函数在频率中间区域随着频率的下降不合理的增加间的相互影响导致.

注意这个泄漏的形式很容易发现：我们仅仅需要在一个小的上升的滤波器宽度序列上比较小波方差估计. 图329右边的图指出在D(4)估计中仍然有少量的泄漏，但是，D(6)和LA(8)很好地跟随彼此的事实表明，一旦我们接触到D(6)滤波器，就不用考虑泄漏.

图333显示了在每个尺度下，伴随着真实小波方差的3个95%置信区间，作为物理尺度的函数小波方差估计值绘的图. 这个图指出这些序列的方差主要是由于在尺度6.4m和更长处的波动，它与海洋内深层射流和内部波动有关. 注意作为 $\log_{10}(\tau_j)$的函数 $\log_{10}(\hat{\nu}_X^2(\tau_j)$在小尺度0.1到0.4m和中等尺度0.8到6.4m上近似线性变化(然而，近似线性的程度是很有限的，在每种情况下，小于尺度的数量阶). 与这两个区域有关的斜率分别为0.7和2.5，小尺度主要由湍流影响，对于作为 $\log_{10}(\tau_j)$的函数 $\log_{10}(\hat{\nu}_X^2(\tau_j))$的斜率表明湍流以与指数 $\alpha=-1.7$ 的幂规律一致的速率重复出现，这个结果可以用作物理模型. 在中间尺度下，指数 $\alpha=-3.5$ 幂规律重复出现可以解释为内部波动和湍流区域之间的一个过渡区域.

图333中的三个置信区间是基于有由 $\hat{\eta}_1$、η_2和 η_3决定的等价自由度 η 的 χ^2 近似. 等价自由度的真实值在表333中列出. 对于七个最小的尺度，从一个实际工作者的观点看，由这三个方法给出的置信区间可互换，但是不要惊讶，在两个最大的尺度下，一致性打破. 置信区间可以用来估计，例如，估计这个特殊序列在尺度25.6m处的波动是否与在海洋其他地方取到的其他组吻合.

最后，让我们再一次说明，小波方差怎样可以轻易地用于处理作为在某些区域局部平稳的最好模型的时间序列. 在5.10节，在图195中描述的，垂直切变

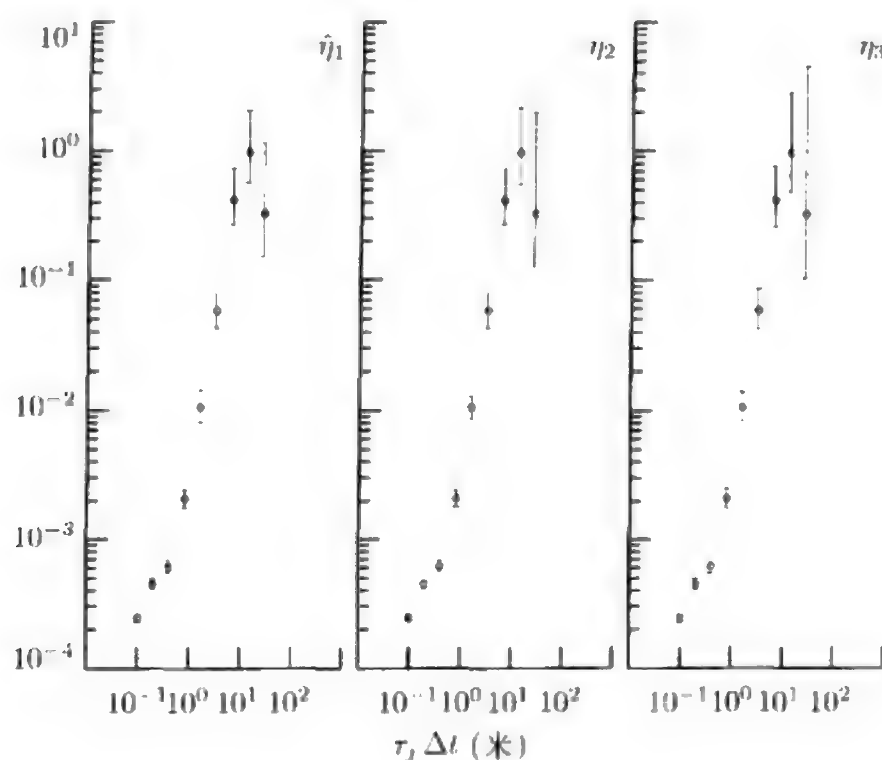


图 333　为垂直海洋切变序列作出的对于 D(6)小波方差的 95%置信区间．这个区间是基于无偏极大重叠离散小波变换估计量(图 329 的＋和上面的 o)和有着由从左到右，式(313d)的$\hat{\eta}_1$；使用由式(313)给出的 $S_X(\cdot)$名义上的模型得到的式(313d)的 η_2；式(314c)的 η_3，所确定的等价自由度的 χ^2 近似分布上(表 333 列出了等价自由度的值)

表　333

	j								
	1	2	3	4	5	6	7	8	9
$\hat{\eta}_1$	1 890	1 027	584	289	94	82	32	20	8
η_2	2 850	1 633	899	359	173	78	31	17	5
η_3	2 046	1 020	508	251	123	59	27	11	3
M_j	4 091	4 081	4 061	4 021	3 941	3 781	3 461	2 821	1 541

注：结合图 333 显示的 D(6)小波方差估计$\hat{\nu}_X^2(\tau_j)$给出的自由度$\hat{\eta}_1$，η_2和 η_3的等价次数(接近最近的整数)．最后一行给出了在每个尺度下的极大重叠离散小波变换小波系数的个数．

序列的 LA(8)多分辨分析的讨论中，我们注意到，在浅的地方，突出的脉冲首先出现在 $\tau_5\Delta t=1.6$m 处，但是随着深度的增加，在较小的尺度下变得更加突出．这里，我们可以通过由式(324)计算的依赖于时间的 LA(8)小波方差证实这个模式(注意用 LA(8)小波而不是 D(6)小波是由于前者较好的时间序列特征)．图 195 中脉冲的大小大约 50m．如果令 $N_S=257$，每段跨度 25.7m，它是一个合

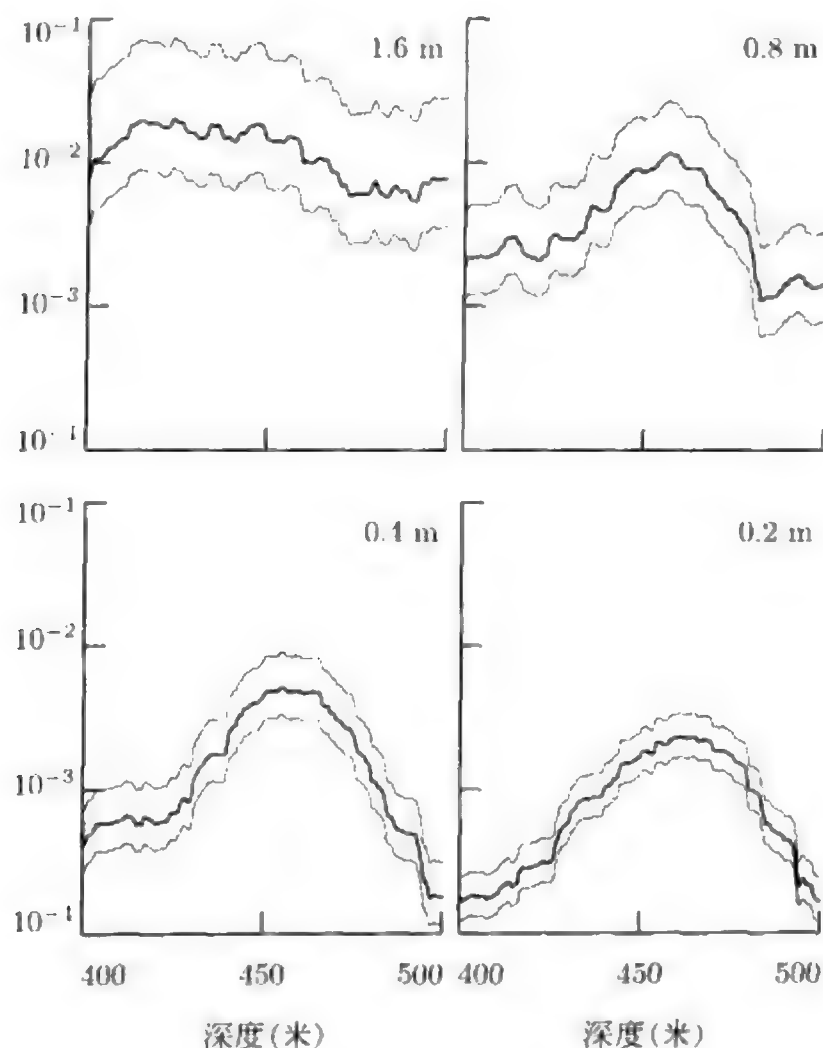


图 334 对于切变数据，估计 LA(8)小波方差，用深度在 450m 左右的 257 个极大重叠离散小波变换小波系数的分段的计算．如图中所标记的，物理尺度是 0.2，0.4，0.8 和 1.6m(即 $\tau_j \Delta t$，$j=2, \cdots, 5$)．估计值上下的细曲线表示按照式(313c)，用 $\eta=257/2^j$ 得到的 95％置信区间(这与用式(314c)的 η_3 符合)

理的区间，在它之上去假设局部平稳性．对于尺度 $\tau_5 \Delta t=1.6$ 低至 $\tau_2 \Delta t=0.2\text{m}$，得到的估计值在图 334 中以粗线绘出，一起的还有通过 $\hat{\nu}^2_{X,t}(\tau_j)$ 分布的 χ^2 近似计算出的高的和低的 95％置信区间(细曲线)(按照式(314)中的 η，等价自由度取值为 $\eta=N_S/2^j$)．注意由于这个图的垂直轴是对数，越过每一个尺度，置信区间的高度看起来与估计方差无关(实际上，高度与估计值成比例)．按科学的方法，重要的观测值是一个关于图 195 中多分辨分析的一个证实，即在小波方差中与脉冲一致的顶点看起来随着尺度的减小深度增加，表明在尺度 1.6m 处的湍流驱使更深层的比较短尺度湍流(参见 Serroukh 等，2000，适合于补充去掉高斯假设的处理)．

8.10 小结

令$\{X_t: t=\cdots, -1, 0, 1, \cdots\}$是一个离散参数实值随机过程，它的 d 阶向后差分

$$Y_t \equiv (1-B)^d X_t \equiv \sum_{k=0}^{d} \binom{d}{k} (-1)^k X_{t-k}$$

是一个具有谱密度函数 $S_Y(\cdot)$和均值 μ_Y 的平稳过程(这里 d 是一个非负整数，μ_Y 不必要为零). 令 $S_X(\cdot)$表示对$\{X_t\}$的谱密度函数，对于它我们有 $S_X(f)=S_Y(f)/\mathcal{D}^d(f)$，其中$\mathcal{D}(f)\equiv 4\sin^2(\pi f)$(如果$\{X_t\}$事实上是一个非平稳过程，$S_X(\cdot)$和$S_Y(\cdot)$之间的关系提供对 $S_X(\cdot)$的一种定义). 给定一个第 j 层极大重叠离散小波变换小波滤波器$\{\tilde{h}_{j,l}: l=0, \cdots, L_j-1\}$，它是基于宽度为 L 的 Daubechies 小波滤波器，定义第 j 层小波系数过程为

$$\overline{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l}, \quad t=\cdots,-1,0,1,\cdots$$

(在上式中，回想 $L_j \equiv (2^j-1)(L-1)+1$). 如果 $L \geqslant 2d$，那么$\{\overline{W}_{j,t}\}$是一个平稳过程，其谱密度函数由 $S_j(f)\equiv \widetilde{\mathcal{H}}_j^{(\mathrm{D})}(f) S_X(f)$给出，其中$\widetilde{\mathcal{H}}_j^{(\mathrm{D})}(\cdot)$是$\{\tilde{h}_{j,l}\}$的平方增益函数. 然后，我们可以在尺度 $\tau_j \equiv 2^{j-1}$上定义$\{X_t\}$的小波方差为

$$\nu_X^2(\tau_j) \equiv \operatorname{var}\{\overline{W}_{j,t}\} = \int_{-1/2}^{1/2} \widetilde{\mathcal{H}}_j^{(\mathrm{D})}(f) S_X(f) \mathrm{d}f$$

(这个方差必定有限且与 t 无关). 我们有

$$\sum_{j=1}^{\infty} \nu_X^2(\tau_j) = \operatorname{var}\{X_t\},$$

其中如果$\{X_t\}$是非平稳的，则 $\operatorname{var}\{X_t\}$是无限的. 总之，$\nu_X^2(\tau_j)$表示在尺度 τ_j 的变化所引起的$\{X_t\}$的整体变化(注意 τ_j 事实上是一个非单位标准尺度，通过与$\{X_t\}$相关的样本区间 Δt 相乘成为实际上有意义的尺度$\tau_j \Delta t$).

假设我们给定一个时间序列，它可以看作是过程$\{X_t\}$的一部分 $X_0, \cdots, X_{N-1}$的实现. 只要 $M_j \equiv N-L_j+1>0$，且或(i)$L>2d$，或(ii)$L=2d$，$\mu_Y=0$(任何一个条件都暗示了 $E\{\overline{W}_{j,t}\}=0$，因此 $\nu_X^2(\tau_j)=E\{\overline{W}_{j,t}^2\}$)，我们可以形成 $\nu_X^2(\tau_j)$的一个无偏估计量

$$\hat{\nu}_X^2(\tau_j) \equiv \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \widetilde{W}_{j,t}^2,$$

其中$\{\widetilde{W}_{j,t}\}$是对这个时间序列的第 j 层极大重叠离散小波变换小波系数：

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N}, \quad t=0,1,\cdots,N-1.$$

注意当 $t=L_j-1,\cdots,N-1$ 时，$\widetilde{W}_{j,t}=\overline{W}_{j,t}$，因而对 t 的这些范围 $E\{\widetilde{W}_{j,t}^2\}=\nu_X^2(\tau_j)$（这个关系对 $t<L_j-1$ 并不必要成立，因此取所有 N 的样本均值，可能的 $\widetilde{W}_{j,t}^2$ 一般来说可能产生 $\nu_X^2(\tau_j)$ 的无偏估计量）．在假设 $\{\overline{W}_{j,t}\}$ 是一个高斯过程的条件下，那么对于大的 M_j，随机变量 $\hat{\nu}_X^2(\tau_j)$ 服从均值为 $\nu_X^2(\tau_j)$、方差为 $2A_j/M_j$ 的近似高斯分布，其中

$$A_j=\int_{-1/2}^{1/2}S_j^2(f)\mathrm{d}f$$

（这个结论在 A_j 有限且 $S_j(f)>0$ 几乎处处成立的条件下成立）．基于 $\hat{\nu}_X^2(\tau_j)$，我们用这个结果可以间接地为 $\nu_X^2(\tau_j)$ 构造一个置信区间（参见式(311)），但是，得到的区间可以有一个负的下限，这很不方便，因为在实际中，我们常常用 log/log 尺度来为小波方差估计和它们相应的置信区间绘图．

对 $\nu_X^2(\tau_j)$ 产生置信区间的一个更吸引人的（而渐近等价的）途径是假设 $\eta\hat{\nu}_X^2(\tau_j)/\nu_X^2(\tau_j)$ 与 χ_η^2 有着相同的分布，即一个具有 η 为等价自由度的 χ^2 随机变量．这个假设导致如下形式的置信区间：

$$\left[\frac{\eta\hat{\nu}_X^2(\tau_j)}{Q_\eta(1-p)},\frac{\eta\hat{\nu}_X^2(\tau_j)}{Q_\eta(p)}\right],$$

其中 $Q_\eta(p)$ 是对 χ_η^2 分布的 $p\times100\%$ 百分率点．以上结论需要 η 的值，它可以使用下述三种方法之一确定．

[1]我们可以使用

$$\hat{\eta}_1\equiv\frac{M_j\hat{\nu}_X^4(\tau_j)}{\hat{A}_j}$$

估计 η，其中

$$\hat{A}_j\equiv\frac{1}{2}\int_{-1/2}^{1/2}[\hat{S}_j^{(\mathrm{p})}(f)]^2\mathrm{d}f=\frac{\hat{\nu}_X^4(\tau_j)}{2}+\sum_{\tau=1}^{M_j-1}(\hat{s}_{j,\tau}^{(\mathrm{p})})^2;$$

这里 $\{\hat{s}_{j,\tau}^{(\mathrm{p})}\}$ 是对 $\{\widetilde{W}_{j,t}\}$ 的自协方差序列的偏差估计量，即

$$\hat{s}_{j,\tau}^{(\mathrm{p})}\equiv\frac{1}{M_j}\sum_{t=L_j-1}^{N-1-|\tau|}\widetilde{W}_{j,t}\widetilde{W}_{j,t+|\tau|},\quad 0\leqslant|\tau|\leqslant M_j-1.$$

[2]如果我们假设 $S_X(\cdot)=aC_X(f)$，其中 $C_X(\cdot)$ 是一个已知函数，并且 a 是一个未知的倍增常数，那么我们可以对 $\{\overline{W}_{j,t}\}$ 的谱密度函数表示为

$$S_j(f)=a\,\widetilde{\mathcal{H}}_j^{(\mathrm{D})}(f)C_X(f)\equiv aC_j(f),$$

并且置 η 为

$$\eta_2\equiv\frac{2\left(\sum_{k=1}^{\lfloor(M_j-1)/2\rfloor}C_j(f_k)\right)^2}{\sum_{k=1}^{\lfloor(M_j-1)/2\rfloor}C_j^2(f_k)}.$$

[3]使用要求极大重叠离散小波变换小波滤波器 $\{\tilde{h}_{j,l}\}$ 的带通性质的论证，我们可以置 η 为

$$\eta_3=\max\{M_j/2^j,1\}.$$

给 η 取值的三种方法的相应的矩阵在 8.4 节的最后讨论.

最后，注意由于第 j 层极大重叠离散小波变换小波滤波器的带通特性，我们可以把在尺度 τ_j 上的小波方差的估计转化成在滤波器通带上的谱密度函数 $S_X(\cdot)$的一个估计. 这个方法使用 $\hat{C}_j \equiv 2^j \hat{\nu}_X^2(\tau_j)\Delta t$ 估计在倍频程频带 $\left[\frac{1}{2^{j+1}\Delta t}, \frac{1}{2^j\Delta t}\right]$上的谱密度函数. 得到的估计量是一个分段常数.

8.11 练习

[8.1] 令$\{\epsilon_t\}$是一个有着零均值和方差 σ_ϵ^2 的白噪声过程(由定义，这个过程有由式(268d)给出的自协方差序列). 注意式(267b)可以用来证明，对于一个白噪声过程，对于所有的 f，谱密度函数由 $S_\epsilon(f)=\sigma_\epsilon^2$ 给出(为了方便假设 Δt 是 1). 导出$\{\epsilon_t\}$的小波方差 $\nu_\epsilon^2(\tau_j)$的一个清晰的表达式. 用这个表达式，验证对于一个白噪声过程，式(296d)成立(即所有可能的二进尺度上的小波方差的和等于过程方差 σ_ϵ^2).

[8.2] 使用一个高斯随机数生成器，创建一个有着零均值和单位方差的长为 $N=1\ 024$ 的白噪声过程的实现. 使用式(306b)的无偏极大重叠离散小波变换估计量$\hat{\nu}_\epsilon^2(\tau_j)$估计小波方差$\nu_\epsilon^2(\tau_j)$，$\tau_j=1, 2, 4, \cdots, 512$，对于生成的序列用下面三个小波滤波器中的任意两个：哈尔、D(4)或 LA(8). 重复使用式(306c)的有偏极大重叠离散小波变换估计量. 在 log/log 尺度下，绘出 τ_j 的函数的 $\nu_\epsilon^2(\tau_j)$估计图. 借助于一个幂规律过程解释你的图.

[8.3] 假设$\{X_t\}$是一个如式(288a)定义的随机游动过程. 证明，对于哈尔小波，我们有

$$\nu_X^2(\tau_j) = \frac{\sigma_\epsilon^2}{6}\left(\tau_j + \frac{1}{2\tau_j}\right),$$

其中 σ_ϵ^2 是用于创建$\{X_t\}$的白噪声过程的方差. 假定 σ_ϵ^2 是 1，在一个 log/log 尺度下，对$j=1, \cdots, 10$，绘出作为 τ_j 函数的 $\nu_X^2(\tau_j)$的图. 怎样可以使你的图对应于由练习[288a]最后部分的答案提出的结果?

[8.4] 使用一个高斯随机数生成器，创建一个长为 $N=1\ 024$ 的随机游动过程$\{X_t\}$的实现(见前一个练习). 用无偏和有偏极大重叠离散小波变换估计量$\hat{\nu}_X^2(\tau_j)$和$\tilde{\nu}_X^2(\tau_j)$(式(306b)和(306c))估计小波方差 $\nu_X^2(\tau_j)$，$\tau_j=1, 2, 4, \cdots, 512$，对于生成的序列，用下面三个小波滤波器中的任意两个：哈尔、D(4)或 LA(8). 在 log/log 尺度下以绘出你的作为 τ_j 的函数的 $\nu_X^2(\tau_j)$估计. 借助于一个幂规律过程解释你的图.

[8.5] 证明，如果$\{X_t\}$是一个平稳过程，小波方差的无偏和有偏极大重叠离散小波变换估计量(分别是式(306b)和式(306c))之间预期的差异随着 $N\to\infty$趋于 0. 如果$\{X_t\}$是一个有着平稳一阶向后差分的非平稳过程，相同的结果也成立?

[8.6] 假设$\{X_t\}$是有着零均值的一阶平稳自回归过程，即我们可以写 $X_t=\phi X_{t-1}+\epsilon_t$，其中 $|\phi|<1$，且$\{\epsilon_t\}$是一个零均值白噪声过程. $\{X_t\}$的自

协方差序列$\{s_\tau\}$由$s_{X,\tau}=\phi^{|\tau|}s_{X,0}$给出，其中$s_{X,0}$是过程的方差. 像式(296a)和(306a)定义且假定利用哈尔小波，令$\{\overline{W}_{1,t}\}$是单位尺度极大重叠离散小波变换小波系数过程，且令$\{\widetilde{W}_{1,t}\}$是基于时间序列X_0，…，X_{N-1}上的单位尺度极大重叠离散小波变换小波系数. t在什么范围内我们有$\widetilde{W}_{1,t}=\overline{W}_{1,t}$? 对于所有的$t$确定$E\{\widetilde{W}^2_{1,t}\}$. 在什么情况下我们有$E\{\widetilde{W}^2_{1,0}\}=\nu^2_X(\tau_1)$?

[8.7] 令X_0，…，X_{N-1}是过程$\{X_t\}$的一部分，这个过程的d阶向后差分形成一个平稳过程. 令$M>N$是被2^j整除的2的任一次幂，且如式(308c)构造X'_t. 令$W'_{j,t}$表示基于宽$L\geqslant 2d$的一个Daubechies小波滤波器上的$\{X'_t\}$的第j层离散小波变换小波系数，令L'_j如式(146a)定义. 在假设$\left\lfloor\frac{N}{2^j}-1\right\rfloor\geqslant L'_j$下，证明，在被$2^{j/2}$整除的前提下，式(308d)所显示的小波系数子序列等于从无穷序列$\{\overline{W}_{j,t}\}$选出的随机变量(这在式(296a)中定义). 在创建X'_t中，如果我们用X_0，…，X_{N-1}的样本均值填充X_t而不是用0将会有什么变化?

[8.8] 假设$\{X_t\}$是使单位尺度极大重叠离散小波变换小波系数过程$\{\overline{W}_{1,t}\}$是平稳的，且具有谱密度函数$S_1(\cdot)$和平方可和的自协方差序列$\{s_{1,\tau}\}$. 定义相应的单位尺度离散小波变换小波系数过程为$\{2^{1/2}\overline{W}_{1,2t+1}: t=\cdots, -1, 0, 1, \cdots\}$(参见式(70c)). 证明这个过程是平稳的，并且有一个对于$|f|\leqslant 1/2$与$S_1\left(\frac{f}{2}\right)+S_1\left(\frac{f}{2}+\frac{1}{2}\right)$成比例的谱密度函数(提示：回想练习[23b]).

[8.9] 验证式(309).

[8.10] 假定我们再写式(314a)的逼近包含在奈奎斯特频率的项，那时有

$$\hat{\nu}^2_X(\tau_j)\approx\frac{2}{M_j}\sum_{k=1}^{\lfloor(M_j-1)/2\rfloor}\hat{S}^{(\mathrm{p})}_j(f_k)+\frac{1}{M_j}\hat{S}^{(\mathrm{p})}_j\left(\frac{1}{2}\right)\mathcal{I}_{M_j},$$

其中当M_j是偶数时$\mathcal{I}_{M_j}$为单位，当M_j时奇数时为零，现在仅仅认为近似$\hat{S}^{(\mathrm{p})}_j(0)/M_j=\overline{W}^2_j$可以忽略. 大样本理论提出$\hat{S}^{(\mathrm{p})}_j\left(\frac{1}{2}\right)/S_j\left(\frac{1}{2}\right)$有一个$\chi^2_1$分布并且这个随机变量与上面求和式中的所有随机变量无关. 在假设

$$\frac{\eta'_2\,\hat{\nu}^2_X(\tau_j)}{\nu^2_X(\tau_j)}\overset{\mathrm{d}}{=}\chi^2_{\eta'_2}$$

下，使用一个等价自由度的论证证明

$$\eta'_2=\frac{\left(2\sum_{k=1}^{\lfloor(M_j-1)/2\rfloor}C_j(f_k)+C_j\left(\frac{1}{2}\right)\mathcal{I}_{M_j}\right)^2}{2\sum_{k=1}^{\lfloor(M_j-1)/2\rfloor}C^2_j(f_k)+C^2_j\left(\frac{1}{2}\right)\mathcal{I}_{M_j}}$$

(参见式(314b)中的η_2).

[8.11] 基于式(313c)(η由式(313d)中的$\hat{\eta}_1$估计)和基于D(4)小波方差$\nu^2_X(\tau_1)$的D(4)极大重叠离散小波变换无偏估计量，对在图42的标题中列出的16点时间序列$\{X_{1,t}\}$的$\nu^2_X(\tau_1)$计算一个95%置信区间. 对于$\nu^2_X(\tau_1)$计算第二个95%置信区间，但这次用式(314c)中的等价自由度η_3. 你认为哪

一个等价自由度的大小用在这里更合适?(注意，表 263 给出了这里可能用到的 χ^2 分布的一些百分率点).

[8.12] 证明，在假设过程均值是未知的因此用样本均值估计的条件下，式(322)的对 τ_1 的 Allan 方差与 2 样本的样本方差成比例. 因此，Allan 方差有时叫做"2 样本"方差，它就是记号 $\sigma_Y^2(2, \tau_j)$ 中"2"的根本理由.

[8.13] 证明，如果 $\{\bar{Y}_t(\tau_1)\}$ 是一个有着方差 σ_Y^2 的白噪声过程，那么我们有 $\sigma_Y^2(2, \tau_1)=\sigma_Y^2$，其中 $\sigma_Y^2(2, \tau_1)$ 是式(322)的 Allan 方差(在白噪声的情形下 Allan 方差简化为过程方差，这是式(322)中因子 1/2 的根本理由，没有它我们不能得到这样的一致).

第 9 章

长记忆过程的分析与综合

9.0 引言

在第 8 章我们注意到，作为尺度 τ_j 的函数的小波方差 $\nu_X^2(\tau_j)$ 的估计的 log/log 图形能帮助我们识别时间序列，该序列能很好地模拟为幂规律过程 $\{X_t\}$；即一过程的谱密度函数 $S_X(\cdot)$ 使 $S_X(f) \propto |f|^{\alpha}$. 当 $\alpha<0$，$f\to 0$ 时，$S_X(f)\to\infty$，我们就说 $\{X_t\}$ 呈现“长记忆”(参见 7.6 节中的精确定义和评论). 在微观和宏观现象的科学研究工作中广泛地发现和研究这种过程：

- 通过细胞膜的电压波动(Holden, 1976).
- 沙粒通过沙漏一个小时的密度变化(Schick and Verveen, 1974).
- 高速公路的交通流量变化(Musha and Higuchi, 1976).
- 地球物理学中钻孔的阻抗变化(Walden, 1994).
- 地球旋转变化(Munk and MacDonald, 1975).
- 星系 X-射线的时间变化(McHardy and Czerny, 1987).

对于长记忆过程，随着 τ_j 变大，作为 $\log(\tau_j)$ 函数的 $\log(\nu_X^2(\tau_j))$ 图表现出斜率为 $-\alpha-1$ 的线性变化. 这个简单的关系暗示小波特别善于处理这种过程. 在最近几年里，许多人，例如，Flandrin(1992)，Masry(1993) 和 Wornell(1993) 等意识到小波变换能较好地研究长记忆相关过程和相关的过程(引用的参考文献主要集中于连续时间非平稳过程的小波级数表示).

本章我们将考虑离散小波变换用于平稳长记忆过程和相关非平稳过程的分析和综合的方法，在 log-频率/log-幂坐标轴中，随着 $f\to 0$，至少在频率 f 的若干倍频程，这两种过程的谱密度函数的图形近似于具有负斜率的直线. 在 9.1 节，我们将讨论离散小波变换近似去除分形差分(FD)过程的相关性(这些过程给出了一些具有长记忆特征的简单易处理模型，参见 7.6 节关于分形差分过程基本性质的讨论). 去相关的属性是基本的：鉴于分形差分过程中的随机变量有高度的相

关性，离散小波变换产生一些新的随机变量族，即小波系数，(在尺度内及之间)它近似地去掉相关性，因此更适合于统计分析. 这个基本结果常用于两个方面. 首先，通过随机产生离散小波变换系数从而利用离散小波逆变换组合这些数据产生输出过程的方法，能够近似模拟长记忆过程(参见 9.2 节). 其次，给定意欲模拟化为分形差分过程的时间序列，我们可以使用近似极大似然估计(MLE)来估计该过程中的两个未知参数，极大似然估计直接以离散小波变换的去相关性为基础，从而最大化简单似然函数(9.3 节和 9.4 节). 基于著名的最佳性质而广泛地把极大似然估计用于统计时，通过小波方差和尺度的最小二乘拟合(log/log 空间)，也可以估计分形差分过程的未知量(9.5 节). 最小二乘估计虽然可论证是次佳的，但是容易计算. 在 9.6 节我们将考虑去相关性质的另一种用途. 在平稳分形差分过程的零假设前提下，对于有用的方差的齐性，设计了一个简单的实验. 在小结(9.9 节)以前，使用原子钟和尼罗河级数(9.7 节和 9.8 节)来说明此章中所讨论的方法.

9.1 长记忆过程的离散小波变换

让我们开始用一个例子研究长记忆过程的离散小波变换的统计属性. 图 342 的左下图显示了具有零均值和参数 $\delta=0.4$ 与 $\sigma_\varepsilon^2=1.0$ 的平稳分形差分过程 $\{X_t\}$ 的长度为 $N=1\,024$ 的实现 $\mathbf{X}$. 该模拟序列用 Davies-Harte 方法(7.8 节描述)产生，并表现了典型长记忆过程的大尺度变量. 为了量化这些变量，我们指出样本自相关序列去延迟在右下图的 $\tau=32$(自相关序列中的值用描绘与零偏差的垂线来表示). 这个序列定义为

$$\hat{\rho}_{X,\tau} \equiv \frac{\sum_{t=0}^{N-1-\tau} X_t X_{t+\tau}}{\sum_{t=0}^{N-1} X_t^2}, \quad \tau = 0,1,\cdots,N-1,$$

该序列是对 $\rho_{X,\tau}\equiv s_{X,\tau}/s_{X,0}$ 的估计，其中 $\{s_{X,\tau}\}$ 是平稳分形差分过程的自协方差序列(ACVS)(有关自协方差序列的描述参见 7.4 节). 注意已估计的自相关序列在所有展示的延迟上表现出实质的正自相关性，这也是典型的长记忆过程.

上述时间序列，我们显示它的基于 LA(8)小波，在 $J_0=7$ 层的部分离散小波变换. 小波函数和尺度系数 $\mathbf{W}_1$，…，$\mathbf{W}_7$ 和 $\mathbf{V}_7$ 循环平移，然后作为时间的函数对 LA(8)滤波器的相位性质绘出系数的图形(参见 4.8 节). 小波系数的图描述了两个重要特征. 第一，这些系数的平均值随尺度增加. 第二，尽管模拟序列本身显示出长项相关性的强可视迹象，但是在已给定尺度内的小波系数之间没有明显的模式(除可论证的 $\mathbf{W}_7$ 外，$\mathbf{W}_7$ 的系数符号交替变化，但仅仅从 8 个系数中推出很多结果是有争议的). 我们可以通对循环平移的小波系数的每个向量通过计算样本自相关序列来量化这种自相关性的不足. 在图 342 的右列绘出了样本自相关

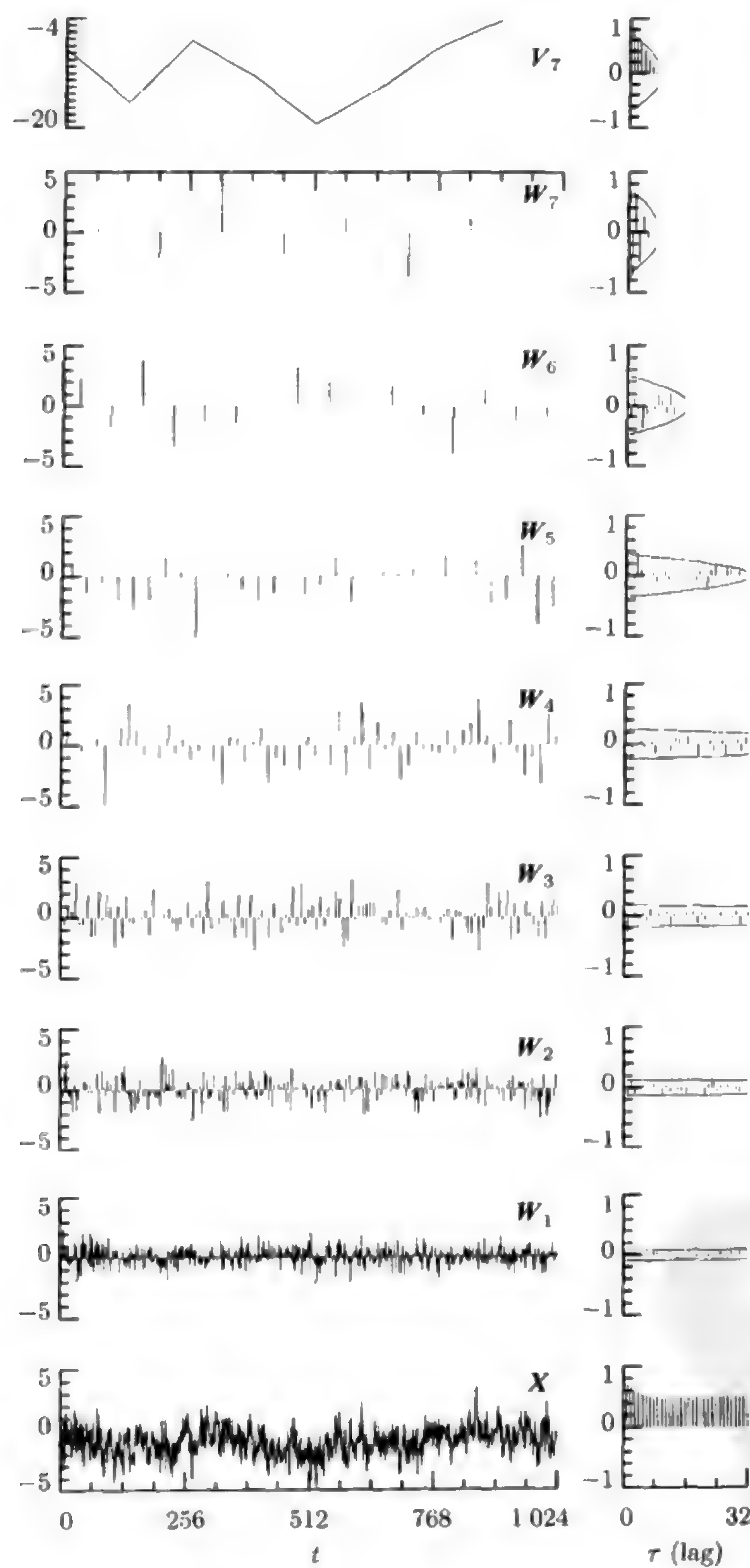


图 342 对于模拟 FD(0.4)时间序列的 LA(8)离散小波变换系数和样本自相关序列

序列(与零的偏差)的图. 在样本自相关序列中的值总体来说比 $\mathbf{X}$ 本身数量级上要小得多，但我们看到平均值随尺度增加. 正如下面讨论所描述的，这种增加能够归因于抽样的变化性. 随着尺度的增加，$\mathbf{W}_j$ 中系数的个数 $N_j \equiv N/2^j$ 将减少. 如果 $\mathbf{W}_j$ 中的值实际上是高斯白噪声的实现，则统计理论指出对于第 j 层尺度样本自相关序列的近似量 95%应下降到在界限 $\pm\frac{2\sqrt{(N_j-\tau)}}{N_j}$ 之间(见每个图的上面和下面的曲线，更细节的理论，参见 Fuller，1996，推论 6.3，6.2). 注意这些界限随 N_j 的减少而发散，在界限外的样本自相关序列值总数和归于抽样变化的那些样本自相关序列值总数和是一致的.

现在让我们验证图 342 所显示的小波系数值增加的比例和 FD(0.4)过程的已知属性是一致的. 在 4.6 节讨论了用等价滤波器$\{h_{j,l}\}$形成关于尺度 $\tau_j=2^{j-1}$ 的小波系数 $\mathbf{W}_j$，$\{h_{j,l}\}$ 近似地是一个带通滤波器，标称的通带由$[-1/2^j, -1/2^{j+1}]\cup[1/2^{j+1}, 1/2^j]$给出. 如果我们回忆，由于$\{h_{j,l}\}$有单位能量，对于$\{h_{j,l}\}$的平方增益函数$\mathcal{H}_j(\cdot)$必须积分为 1，我们能够产生粗糙的近似

$$\mathcal{H}_j(f) \approx \begin{cases} 2^j & 1/2^{j+1} \leqslant |f| \leqslant 1/2^j; \\ 0 & \text{其他}. \end{cases}$$

若把$\{h_{j,l}\}$用于滤波器$\{X_t\}$，就得到带有谱密度函数的一个过程，谱密度函数由$\mathcal{H}_j(f)S_X(f)$给出(参见式(268b)). 这个过程与 $\mathbf{W}_j$ 中的"循环无约束"系数 $W_{j,t}$ 有关，即在给定尺度上，$W_{j,t}$中除了边界的所有系数不多于 $L-2$ 个(参见 4.11 节，通常 L 是单位尺度小波滤波器的宽度)，若略去这些边界系数，$W_{j,t}$的方差由式(268c)给出，即

$$\operatorname{var}\{W_{j,t}\} = \int_{-1/2}^{1/2} \mathcal{H}_j(f)S_X(f)\,\mathrm{d}f \tag{343a}$$

$$\approx 2\int_{1/2^{j+1}}^{1/2^j} 2^j S_X(f)\,\mathrm{d}f = \frac{1}{\frac{1}{2^j}-\frac{1}{2^{j+1}}}\int_{1/2^{j+1}}^{1/2^j} S_X(f)\,\mathrm{d}f \equiv C_j. \tag{343b}$$

如果我们回忆

$$\frac{1}{b-a}\int_a^b x(u)\,\mathrm{d}u$$

是函数 $x(\cdot)$在区间$[a, b]$上的平均值，则能把 C_j 看作是谱密度函数在倍频程带$\left[\frac{1}{2^{j+1}}, \frac{1}{2^j}\right]$上的平均值.

现在让我们特殊化分形差分过程，此过程具有式(284a)给出的谱密度函数 $S_X(\cdot)$，因此产生

$$C_j = 2^{j+1}\int_{1/2^{j+1}}^{1/2^j} \frac{\sigma_\varepsilon^2}{[4\sin^2(\pi f)]^\delta}\,\mathrm{d}f. \tag{343c}$$

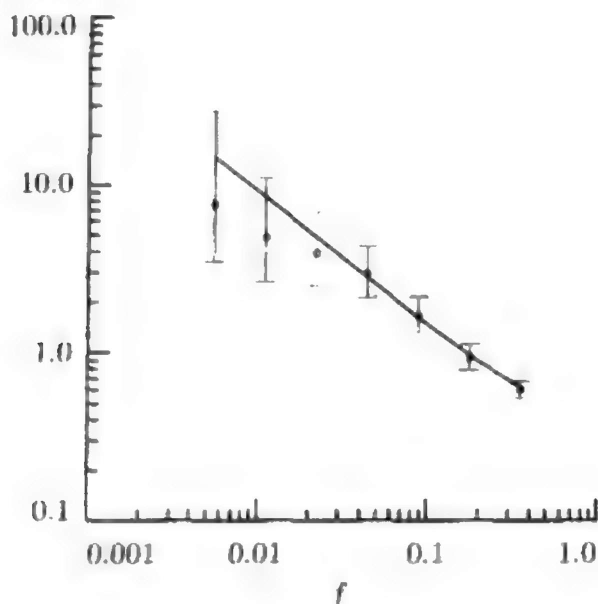


图 344　由图 342 中 LA(8)小波系数的样本方差，从右至左对于层 $j=1,\cdots,7$(圆点)，而 FD(0.4)谱密度函数在倍频程带 $\left[\frac{1}{2^{j+1}},\frac{1}{2^j}\right]$ 的中心频率 $1/2^{j+\frac{1}{2}}$ 取值

因为近似式 $\sin^{2\delta}(\pi f)\approx(\pi f)^{2\delta}$ 在 $f\leqslant\frac{1}{8}$ 时是相当精确的(参见图 282)，对于 $j\geqslant3$，我们可以使用

$$C_j\approx2^{j+1}\int_{1/2^{j+1}}^{1/2^j}\frac{\sigma_\varepsilon^2}{(2\pi f)^{2\delta}}\mathrm{d}f=\sigma_\varepsilon^2\left(\frac{2^{j+\frac{1}{2}}}{2\pi}\right)^{2\delta}\frac{2^{2\delta}-2}{2^\delta(2\delta-1)}\equiv\widetilde{C}_j,\quad\delta\neq\frac{1}{2},\tag{344}$$

并且当 $\delta=1/2$ 时，$\widetilde{C}_j=\sigma_\varepsilon^2 2^{j+\frac{1}{2}}\log(2)/(\pi\sqrt{2})$(在下面的评论与扩展第[4]条中考虑更精确的近似，当 $j=1,2$ 时，这种近似经常使用)．注意 $\widetilde{C}_j\propto[1/2^{j+\frac{1}{2}}]^{-2\delta}$ 其中 $1/2^{j+\frac{1}{2}}$ 是倍频程带 $\left[\frac{1}{2^{j+1}},\frac{1}{2^j}\right]$(关于 log 尺度)的中心频率．因为在这个频率的谱密度函数能够认为是倍频程带的平均值 C_j 的一个代表，我们应该发现 $\mathrm{var}\{W_{j,t}\}$ 与 $S_X(1/2^{j+\frac{1}{2}})$ 彼此接近，进而，作为 $\log(1/2^{j+\frac{1}{2}})$ 函数的 $\log(\mathrm{var}\{W_{j,t}\})$ 或 $\log(S_X(1/2^{j+\frac{1}{2}}))$ 的图形应该近似于斜率为 -2δ 的直线．

作为例子，图 344 的圆圈表示小波系数的样本方差 $\widehat{\mathrm{var}}\{W_{j,t}\}\equiv\|\mathbf{W}_j\|^2/N_j$，小波系数在图 342 中出现并在 log/log 轴绘出了作为中心频率 $1/2^{j+\frac{1}{2}}$，$j=1,\cdots,7$ 函数的图形．叠合的细线是作为这些相同频率的函数 $S_X(1/2^{j+\frac{1}{2}})$ 的图形，这些频率是针对参数为 $\delta=0.4$，$\sigma_\varepsilon^2=1.0$ 的 FD(0.4)过程(该曲线在 log/log 图形中明显接近线性并且最小二乘斜率为 -0.79，该值很好地与理论值 $-2\delta=-0.8$ 吻合)．虽然对于 $j=1,2,3,4$(四个最高频率)，$\widehat{\mathrm{var}}\{W_{j,t}\}$ 很好地与谱密度函数值吻合，而对于 $j=5,6,7$ 不吻合．这些样本方差仅仅使用 32，16，8 个值形成，因此我们需要评估抽样变化性的影响．在假设 $\mathbf{W}_j$ 的元素来源于高斯分布的随机抽样情况下，而这个高斯分布有零均值和未知方差 $\mathrm{var}\{W_{j,t}\}$(由于短暂讨论的去相关属性可以认为是

合理逼近)，由此得出随机变量 $N_j\ \widehat{\text{var}}\{W_{j,t}\}/\text{var}\{W_{j,t}\}$ 服从具有 N_j 个自由度的 χ^2 分布(参见 7.2 节). 我们可以引用练习[313b]论证

$$\left[\frac{N_j\ \widehat{\text{var}}\{W_{j,t}\}}{Q_{N_j}(0.975)},\frac{N_j\ \widehat{\text{var}}\{W_{j,t}\}}{Q_{N_j}(0.025)}\right]$$

对于未知的 $\text{var}\{W_{j,t}\}$ 是一个 95%的置信区间. 在图 344 中表示出了这些区间，并且指出当考虑抽样变化性的时候 $\widehat{\text{var}}\{W_{j,t}\}$ 与 $S_X(1/2^{j+\frac{1}{2}})$ 之间的偏差是不明显的(作为 $\log(1/2^{j+\frac{1}{2}})$ 的函数 $\log(\text{var}\{W_{j,t}\})$ 逼近斜率为 -2δ 的直线的暗示，我们可以通过这些变量进行最小二乘拟合来估计 δ，参见 9.5 节).

现在来看各种离散小波变换如何使平稳分形差分过程很好地去相关.

练习[345] 令 $W_{j,t}$ 和 $W_{j',t'}$ 分别为 $\mathbf{W}_j$ 和 $\mathbf{W}_{j'}$ 中的两个任意非边界系数(即系数不受循环性假设影响). 证明

$$\text{cov}\{W_{j,t},W_{j',t'}\}=\sum_{l=0}^{L_j-1}\sum_{l'=0}^{L_{j'}-1}h_{j,l}h_{j',l'}s_{X,2^j(t+1)-l-2^{j'}(t'+1)+l'},\tag{345a}$$

其中，L_j 是滤波器 $\{h_{j,l}\}$ 的宽度. 另外，证明，当 $j=j'$，$t'=t+\tau$ 时，上式简化为

$$\text{cov}\{W_{j,t},W_{j,t+\tau}\}=\sum_{m=-(L_j-1)}^{L_j-1}s_{X,2^j\tau+m}\sum_{l=0}^{L_j-|m|-1}h_{j,l}h_{j,l+|m|}.\tag{345b}$$

◁

我们可以使用式(345b)来计算对于在第 j 层的非边界小波系数的理论自相关序列，即

$$\rho_{j,\tau}\equiv\frac{\text{cov}\{W_{j,t},W_{j,t+\tau}\}}{\text{var}\{W_{j,t}\}},\quad \tau=\cdots,-1,0,1,\cdots$$

(回忆 $\text{var}\{W_{j,t+\tau}\}=\text{var}\{W_{j,t}\}=\text{cov}\{W_{j,t},W_{j,t}\}$). 图 346a 表示了与哈尔、D(4)和LA(8)小波滤波器结合的 FD(0.4)过程在 $j=4$ 层和延迟 $\tau=1, 2, 3, 4$ 时的计算结果(注意垂直的轴的范围是从−0.2 到 0.2). 如图 342，在每个自相关序列中的值绘出与零的偏差(这些值接近于零以至于看不清). 在这 12 个图的每个中，在量值上的最大自相关性发生在单位延迟上(对于 D(4)离散小波变换，最大的是 $\rho_{4,1}\doteq-0.140$)，之后自相关序列对于 $\tau\geqslant2$ 变得很接近于零. 注意对于给定的离散小波变换，$|\rho_{j,1}|$ 随层 j 增加，但是计算显示对于所有三个离散小波变换有 $|\rho_{j,1}-\rho_{4,1}|<0.01(j\geqslant5)$，因此单位延迟相关不会增长得超过 $j=4$. 为了比较起见，图 346b 表示对于 FD(0.4)过程的延迟 $\tau=64$ 的自相关序列. 在图中给出的所有自相关性在量值上大于图 346a 中的. 因此，在给定尺度中，所有

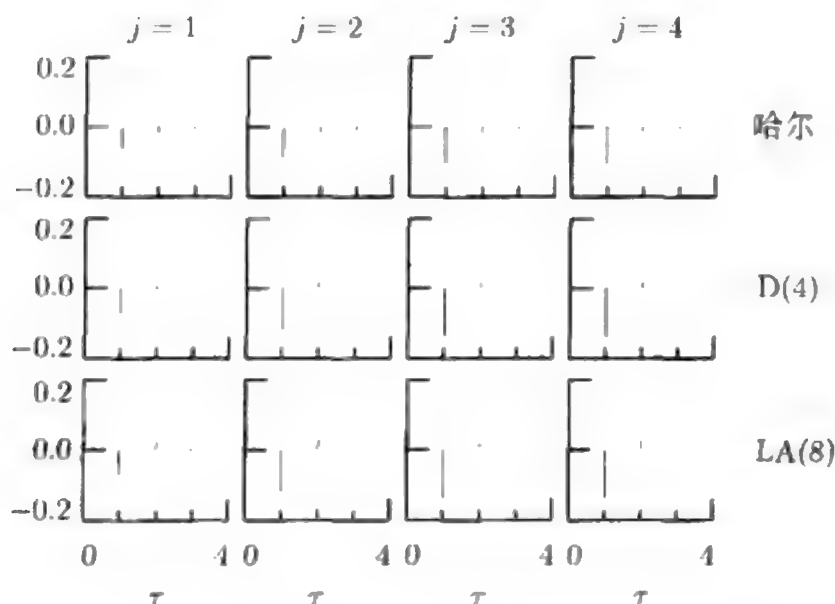


图 346a 对于 FD(0.4)过程中在 $\tau=1, \cdots, 4$ 时的自相关序列哈尔、D(4)和 LA(8)的小波系数. 自相关序列的值从与零的偏差绘起(这些值接近于零以至于看不清)

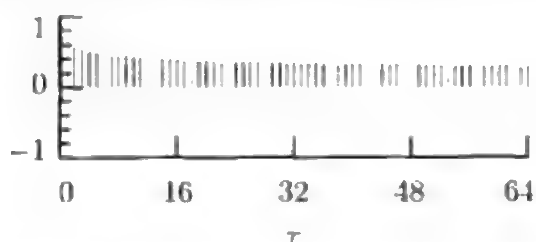


图 346b 对于分形差分过程带有 $\delta=0.4$ 时的自相关序列向外延迟 $\tau=64$

三个离散小波变换的去相关特性都相当好(注意这三个离散小波变换实际上操作相同,尽管——足以感兴趣的——哈尔离散小波变换比其他的稍好点).

为了检验在不同层的小波系数的相关性,对于给定的 j 和 j',我们能用式(345a)来计算在 t 和 t' 的网格上的 $W_{j,t}$ 与 $W_{j',t'}$ 之间的相关性. 图 347a 和图 347b 表示了对哈尔和 LA(8)离散小波变换在 $1\leqslant j<j'\leqslant 4$ 的情况下的计算结果(练习[9.1]建立对于 D(4)的情形下的类似的图). 对于哈尔离散小波变换我们置 $t=2^{|j'-j|-1}$,并且对于 LA(8)离散小波变换,有

$$t=\begin{cases}0, & |j'-j|=1;\\ 1, & |j'-j|=2;\\ 7, & |j'-j|=3,\end{cases} \tag{346}$$

然后令 $t'=t+\tau$,$\tau=-8, \cdots, 8$. 对于 t 和 t' 的这些选择,要确保每个图形在这些指标所有可能的选择下包含极大相关. 通过尺度的相关是小的,但可以看到利用 LA(8)小波比哈尔小波有一个好处:前者的极大绝对相关比后者要小得多.

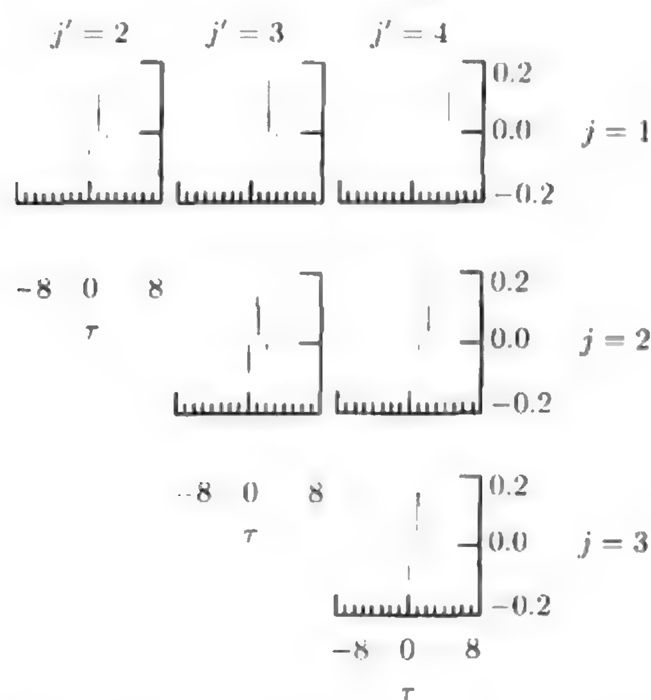


图 347a 由 FD(0.4)过程和对于满足 $1\leqslant j<j'\leqslant 4$ 的层，在哈尔小波系数 $W_{j,t}$ 和 $W_{j',t'}$ 之间的相关性. 设 $t=2^{(j'-j)-1}$ 和 $t'=t+\tau$，而有 $\tau=-8,\cdots,8$，通过所有可能的 t 和 t' 组合，我们捕捉到两系数展示的极大绝对相关性

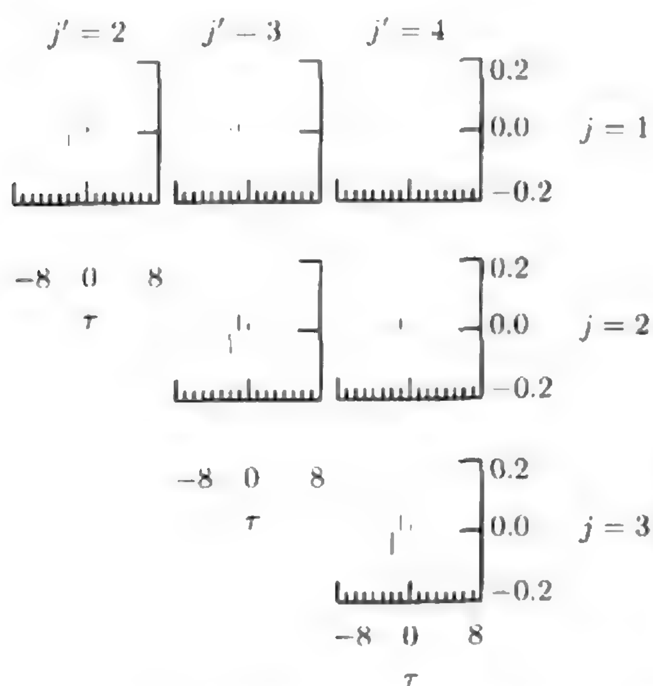


图 347b 如图 347a 中，但现在使用 LA(8)离散小波变换和 t 取值为式(346)

通过公式化等价式(345a)的频域，我们可以得到对离散小波变换的去相关的另一种看法.

练习[348a] 利用自协方差序列是谱密度函数(参见式(267c))的离散小波逆变换的事实，证明

$$\operatorname{cov}\{W_{j,t},W_{j',t'}\}=\int_{-1/2}^{1/2}\mathrm{e}^{\mathrm{i}2\pi f(2^{j'}(t'+1)-2^{j}(t+1))}H_j(f)H_{j'}^{*}(f)S_X(f)\mathrm{d}f,\quad(348\mathrm{a})$$

其中 $H_j(\cdot)$是对应于第 j 层小波滤波器的传递函数. ◁

上式表明对于尺度的相关在数量上 LA(8)比哈尔小波要小得多：如图 107 所示，式(348a)中的传递函数乘积的大小应该随着 L 的增加而变小，这是由于$\{h_{j,l}\}$和$\{h_{j',l}\}$通过带通滤波器后变为更好的近似.

在式(348a)中令 $j=j'$，$t'=t+\tau$，我们得到式(345b)的频域等价式，即

$$\operatorname{cov}\{W_{j,t},W_{j,t+\tau}\}=\int_{-1/2}^{1/2}\mathrm{e}^{\mathrm{i}2\pi 2^{j}f\tau}\,\mathcal{H}_j(f)S_X(f)\mathrm{d}f.\qquad(348\mathrm{b})$$

练习[9.2]使用这个表示式来讨论，当 $S_X(\cdot)$在整个标称通带$[1/2^{j+1},1/2^{j}]$上近似为常数时，对于 $\tau\neq0$，$\operatorname{cov}\{W_{j,t},W_{j,t+\tau}\}\approx0$. 我们基于下面的结果可以得到进一步的了解.

练习[348b] 证明

$$\operatorname{cov}\{W_{j,t},W_{j,t+\tau}\}=\int_{-1/2}^{1/2}\mathrm{e}^{\mathrm{i}2\pi f\tau}S_j(f)\mathrm{d}f,$$

其中

$$S_j(f)\equiv\frac{1}{2^{j}}\sum_{k=0}^{2^{j}-1}\mathcal{H}_j\left(\frac{f+k}{2^{j}}\right)S_X\left(\frac{f+k}{2^{j}}\right).$$ ◁

上面表明，如果我们忽略边界系数，那么 $\mathbf{W}_j$ 中的随机变量可以作为具有谱密度函数 $S_j(\cdot)$的平稳过程的一部分，然而如果 $S_j(\cdot)$接近于一个白噪声过程的谱密度函数. 即如果$S_j(\cdot)$是近似平滑的，这些非边界小波系数将近似无关.

作为一个例子，图 349 显示了一个 FD(0.4)过程(左图)的谱密度函数 $S_X(\cdot)$，标称通带与垂线勾画的离散小波变换系数 $\mathbf{W}_1$，…，$\mathbf{W}_4$ 有关. 右图显示了当使用 LA(8)离散小波变换时的谱密度函数 $S_j(\cdot)$(考虑这些谱密度函数的形状时，回忆图 211 说明的频率反序性). 考虑当 $f\to0$ 时，$S_X(\cdot)$无限地增长，对于非边界小波系数谱密度函数具有最大值和最小值，它们以不超过 3 分贝而区别. 这些 $S_j(\cdot)$没有非常不同的事实可以由任意给定倍频程带时$S_X(\cdot)$改变不剧烈这一事实得到：随着 $S_X(f)\to\infty(f\to0)$倍频程带长度减小，该过程陈述解释了离散小波变换为什么很好地适用于分形差分和其他长记忆过程(有关这一点的进一步讨论见评论与扩展的第[3]条).

像第 4 章，对于一个长为 $N=2^J$ 的时间序列 $\mathbf{X}=[X_0,\cdots,X_{N-1}]^{T}$，让我们把完全离散小波变换写成矩阵形式 $\mathbf{W}=\mathcal{W}\mathbf{X}$(考虑一个完全而不是局部离散小波

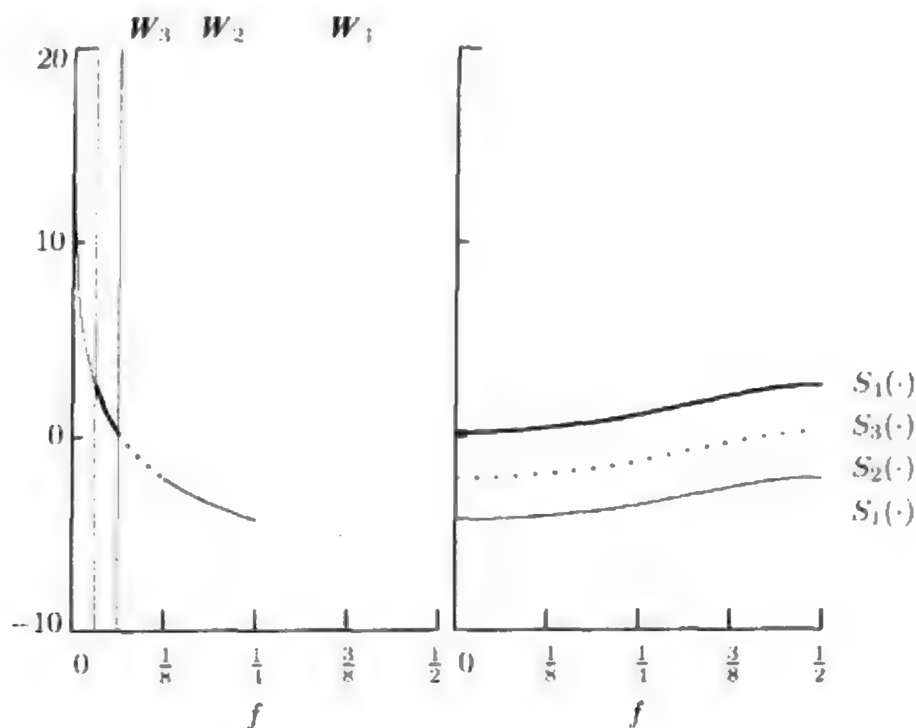


图 349 FD(0.4)过程的谱密度函数(左图)和在 W_1, W_2, W_3, W_4 中非边界 LA(8)小波系数的谱密度函数(右图). 竖轴以分贝为单位,即我们绘出作为 f 的函数 $10\cdot\log_{10}(S_X(f)$的图形). 左图中的垂直线表示四个 W_j 的标称通带

变换的关系是下面评论与扩展第[1]条的主题). $\mathbf{W}$ 的协方差矩阵 Σ_W 为

$$\Sigma_W = \mathcal{W}\,\Sigma_X\,\mathcal{W}^{\mathrm{T}}, \tag{349}$$

其中 Σ_X 是 $\mathbf{X}$ 的协方差矩阵(参见式(262c)). 这个结果是恰当的并且给出了我们希望选择的任何小波滤波器导致的协方差结构,此时适当考虑周期边界条件. 一旦我们形成 $\mathbf{W}$ 的协方差矩阵,就可以计算出相应的相关矩阵:设 $\Sigma_{W,m,n}$ 是 Σ_W 的第(m, n)个元素,则相关矩阵的第(m, n)个元素可以由 $\Sigma_{W,m,n}/\sqrt{(\Sigma_{W,m,m}\Sigma_{W,n,n})}$ 来定义.

图 350、351a 和 351b 给出了由时间序列 $\mathbf{X}$ 生成的 $\mathbf{W}$ 的相关矩阵. $\mathbf{X}$ 是与哈尔、D(4)和 LA(8)离散小波变换联合的 FD(0.4)过程中长为 $N=32$ 的一部分. 通过每个图形中的一系列峰点的线给出了对角元素,所有对角元素是单位元并且常用来衡量非对角相关的大小. 最前峰点和最后峰点分别是元素(31, 31)和(0, 0). 沿着图像的右边缘的黑色节点给出了子空间 $\mathbf{W}_j$ 和 $\mathbf{W}_{j+1}$($j=1$, …, 3)之间的边界. 这些图证实了前面关于各种离散小波变换的去相关属性的结论. 例如,图 350 中对于哈尔离散小波变换远离主对角线的峰点由尺度间的相关所致(参见图 347a). 在图 351b 相应的部分中,对于 LA(8)离散小波变换这些相关是比较小,这与图 347b 一致. 这些图的其余值给了我们一些关于离散小波变换如何处理这些边界小波系数的启示. 如果把关于哈尔离散小波变换(没

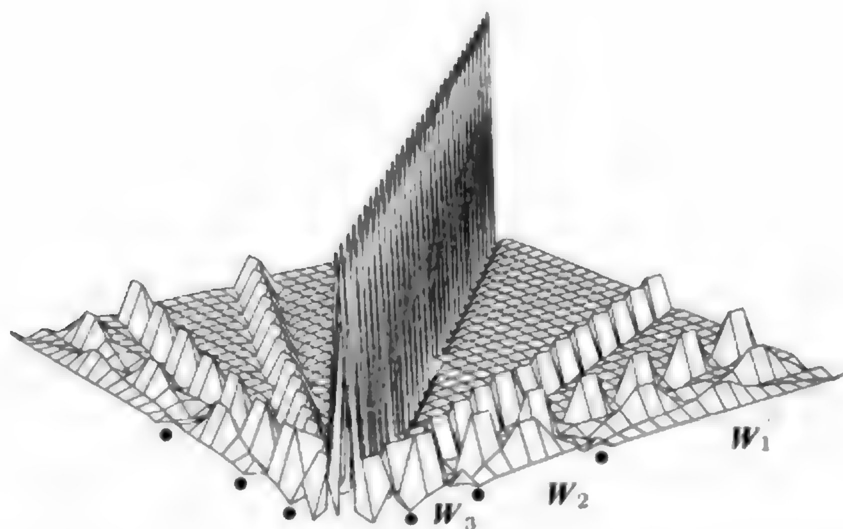


图 350　由分形差分过程中取 $\delta=0.4$ 的长 $N=32$ 时的部分的哈尔小波系数的相关矩阵

有边界系数)和 D(4)离散小波变换(在 $\boldsymbol{W}_1$ 中有一个边界系数，在 $\boldsymbol{W}_2$ 到 $\boldsymbol{W}_5$ 中有两个边界系数)的图形进行比较，可以看到后者涉及边界系数的相关：这些在图像的边缘的峰点，除了在中间近似大小的那些峰点之外. D(4)和LA(8)图的比较显示了以延伸边界为方法，LA(8)离散小波变换减少了影响相关的边界的数量.

最后，让我们回忆在这个例子中，$\boldsymbol{X}$ 有一个相关矩阵，该矩阵的第(m, n)个元素由 $\rho_{X,|m-n|}\equiv s_{X,|m-n|}/s_{X,0}$ 给出，其中 $0\leqslant|m-n|\leqslant 31$. 这个自相关序列在图 346b 中延迟为 64，但是对于延迟小于等于 31，最小值是 $\rho_{X,31}\doteq 0.338$，该值比对 $\boldsymbol{W}$ 的三个已显示相关矩阵的任何非对角元的量要大. 因此，虽然离散小波变换对去掉分形差分过程的相关性不是完美的，但是离散小波变换系数显示的总相关性明显比在 $\boldsymbol{X}$ 中的原始随机变量要小得多.

9.1 节的评论与扩展

[1]让我们考虑图 342 所示的 $J_0=7$ 层的 LA(8)的部分离散小波变换的另一方面. 在右上角，我们表示了对于在尺度系数循环平移向量中的八个值的抽样自相关序列. 在 $\boldsymbol{V}_7$ 中的所有八个值是负的，是在这个特殊模拟序列中的大部分值为负值事实的一个反映(即使分形差分过程$\{X_t\}$有零均值，但是从一个长记忆过程实现的特征化是一个十分长拉伸的趋势，这个拉伸或者以正值为主，或者以负值为主). 在抽样自相关序列中的所有 t 个值均为正，并且有三个在抽样定理所陈述的上限之外. 自认为仅从这个例子做出很多结论是危险的(特别地由于依据抽样定理上限基于假设远多于八个值)；尽管如此，这个例子还是与理论一致，因为考虑到对于分形差分过程的小波系数是近似无关的，尺度系数事实上给出了一个长记忆结构本身. 因此，尽可能处理少的尺度系数是明智的，这就是我们为

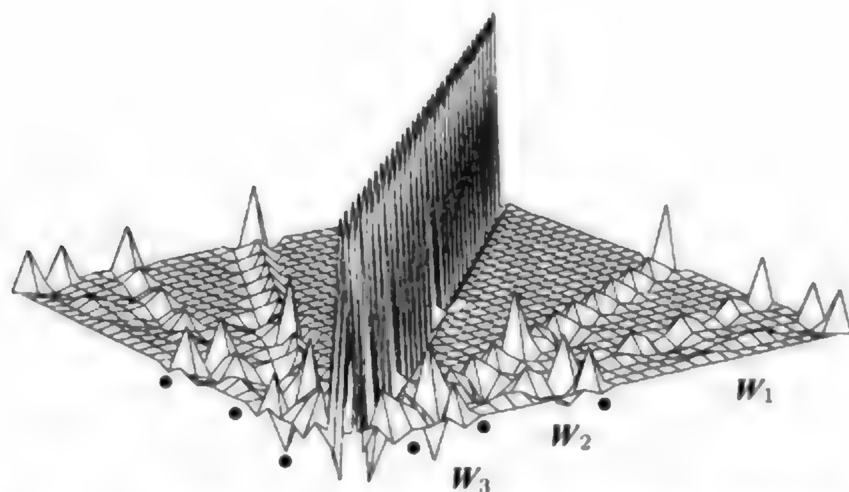


图 351a 如图 350，但是现在使用 D(4)离散小波变换

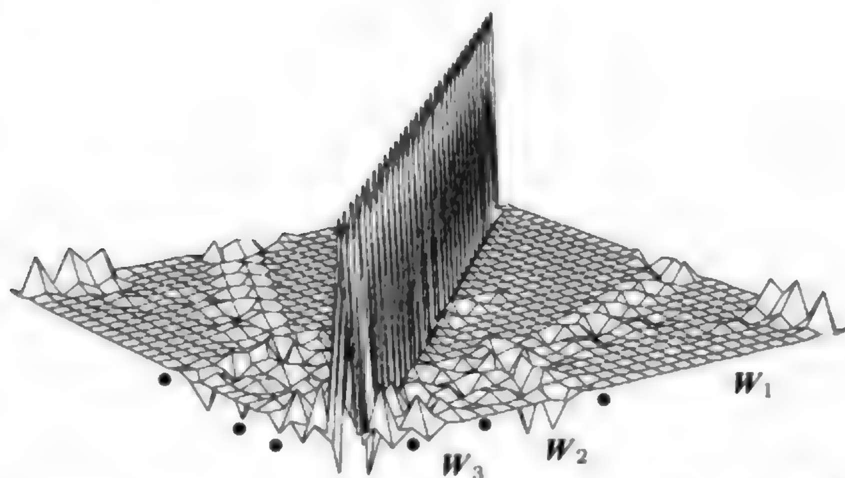


图 351b 如图 350，但是现在使用 LA(8)离散小波变换

什么在这里使用完全离散小波变换而不是部分离散小波变换的原因.

[2]在第 8 章所讨论的概念内容中，注意，图 344 是与小波方差有关的谱密度函数估计的一个例子，并且由式(316)描述，该式借助于式(306b)的极大重叠离散小波变换的无偏差估计$\hat{\nu}_X^2(\tau_j)$；然而，我们这里使用的过程相当于使用式(308b)中基于离散小波变换的有偏差估计量$\tilde{\nu}_X^2(\tau_j)$来代替，因为$2^j\tilde{\nu}_X^2(\tau_j)=\|\mathbf{W}_j\|^2/N_j$.

[3]我们已经讨论了 LA(8)离散小波变换作为一个 FD(0.4)过程的去相关器工作得好(Whitcher，1998，4.1.1 节，证明了，当$-\frac{1}{2}<\delta<\frac{1}{2}$时，可以得到去相关的同样的度，并且哈尔和 D(4)小波也是如此). 这里主要探讨它在其他平稳过程中运行得有多好. 作为一个具体的例子，让我们考虑一个 $p=1$ 阶的自回归过程(AR)，每个式(268e)能写为 $X_t=\phi X_{t-1}+\varepsilon_t$，其中 ϕ 是 AR(1)参数，而$\{\varepsilon_t\}$是一个有零均值和方差 σ_ε^2的白噪声过程. 对于此过程的谱密度函数由式(268f)给出并且在图 352 中绘出在 $\sigma_\varepsilon^2=1$，ϕ 的两个值也就是 0.9(左上图)和-0.9(左下图)

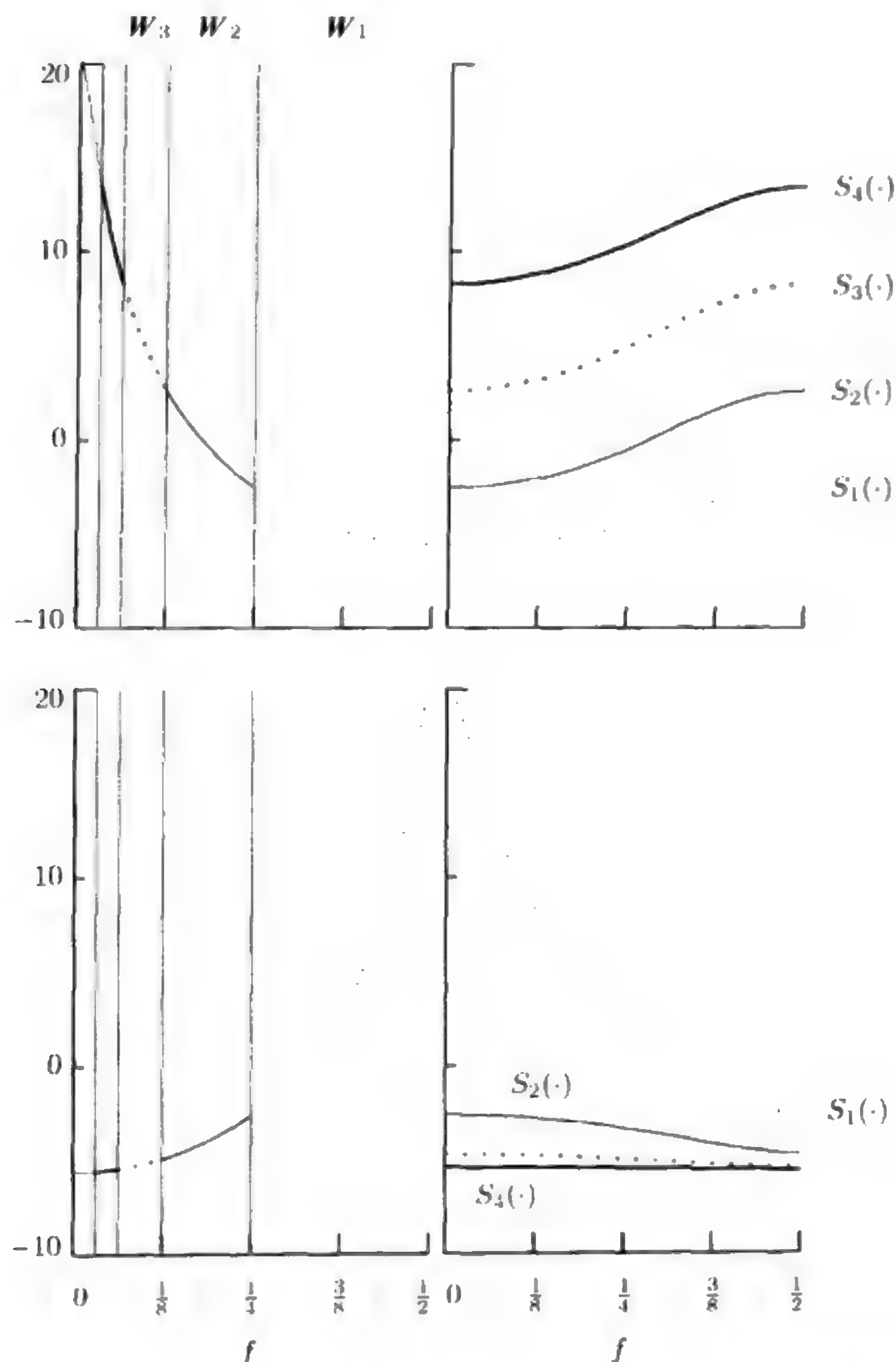


图 352 AR(1)过程中取 $\phi=0.9$(左上图)和 $\phi=-0.9$(左下图)时的谱密度函数，和相应的非边界 LA(8)小波系数在 W_1，…，W_4 时的谱密度函数(右边两图). 纵轴是以分贝为单位的，且左图中的垂直线表示四个 W_j 的标称通带

时的图形. 对于 AR(1)过程，相对应的自相关序列由 $\phi^{|\tau|}$ 给出. 因为这个自相关序列随着 τ 变大指数衰减到零，所以一个 AR(1)过程认为没有长记忆. 图 352 的右部分表示对于非边界小波系数 W_1，…，W_4 的谱密度函数 $S_j(\cdot)$(上图是 $\phi=0.9$，下图是 $\phi=-0.9$). 图 353 表示相对应自相关序列的部分(上部分是 $\phi=0.9$，下部分是 $\phi=-0.9$). 我们看到离散小波变换合理且理想地把 $\phi=0.9$ 过程去

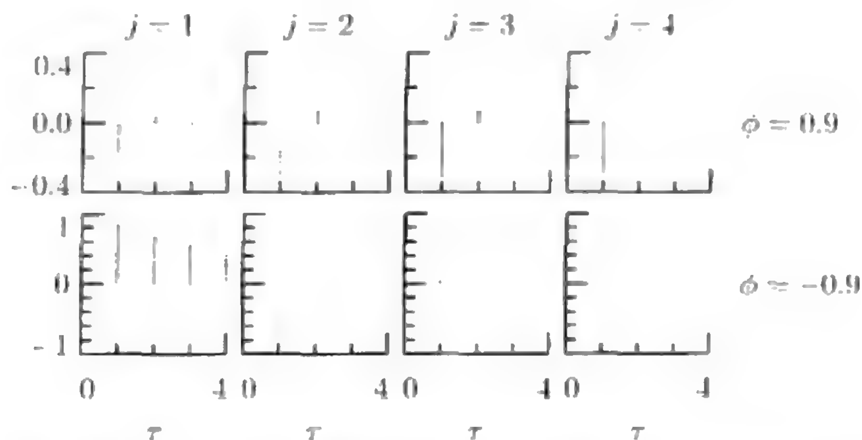


图 353　AR(1)过程在 ϕ 取 0.9(上面一行)和 −0.9(下面一行)时关于 LA(8)小波系数 $W_{j,t}$，$j=1,\cdots,4$ 的自相关序列，$\tau=1,\cdots,4$

相关(但是对于 FD(0.4)过程是不好的，参见图 346a 的下面图形，并且注意到图 352 对应的谱密度函数 $S_j(\cdot)$比图 349 的表现出更多的偏差)；然而，离散小波变换失败于作为 $\phi=-0.9$ 过程的去相关，这是由于 $\boldsymbol{W}_1$ 所表现的自相关序列值不接近于零. 在这里离散小波变换拙劣的原因是对于具有 $\phi=-0.9$ 的 AR(1)过程的谱密度函数在整个标称通带$\left[\frac{1}{4},\ \frac{1}{2}\right]$有相当大的变化. 与 $\boldsymbol{W}_1$ 有关，这引起了一个典型的非平滑 $S_1(\cdot)$. 因此，当离散小波变换对于许多时间序列而作为去相关器使用时，就可能失败，无论何时在已给定的倍频程带有太大的变化. (Whitcher(1998)的 4.1.2 节讨论了覆盖整个范围 $-1<\phi<1$ 的自回归过程. 对于另一类平稳过程，练习[9.3]给读者改变产生绘出的与图 352 和图 353 类似的图形.)

当离散小波变换在此不成功时，注意到对 $\phi=\pm 0.9$ 的谱密度函数是彼此镜像的，以至于如果用通常的方法能在相反方向上做出$\left[0,\ \frac{1}{2}\right]$的一个倍频程带分裂，那么可以把对 $\phi=-0.9$ 的谱密度函数分成与当 $\phi=0.9$ 时变化一样小的区间. 事实上，这个反转可以通过选择规范正交变换来完成，该变换是在小波包树范围内的离散小波变换的“镜像”(参见 6.1 节). 更一般地，在时间序列基于小波的仿真的内容中，对于适应地去掉给定的时间序列的相关性，Percival 等(2000b)讨论了一个从小波包树中适当地选取变换的方案.

[4]这里讨论 C_1 与 C_2 的近似，这个近似比用式(344)得到的更精确(这个近似在所要求的 $j\geqslant 3$ 时也可以使用). 思想是用在积分区间$[1/2^{j+1},\ 1/2^j]$的中点的泰勒(Taylor)级数展开来近似式(343c)的被积函数. 令 $x=\pi f$，$\beta=2\delta$. 则有

$$C_j = 2^{j+1}\int_{1/2^{j+1}}^{1/2^j}\frac{\sigma_\varepsilon^2}{[4\sin^2(\pi f)]^\delta}\mathrm{d}f \equiv \frac{2^{j+1}\sigma_\varepsilon^2}{\pi 2^\beta}\int_{\pi/2^{j+1}}^{\pi/2^j}\frac{u(x)}{x^\beta}\mathrm{d}x,$$

其中 $u(x)\equiv x^\beta/\sin^\beta(x)$. $u(\cdot)$的一阶导数和二阶导数分别由

$$u'(x)=\frac{\beta x^{\beta-1}}{\sin^{\beta}(x)}-\frac{\beta x^{\beta}\cos(x)}{\sin^{\beta+1}(x)}$$

和

$$u''(x)=\frac{\beta x^{\beta}+\beta(\beta-1)x^{\beta-2}}{\sin^{\beta}(x)}-\frac{2\beta^{2}x^{\beta-1}\cos(x)}{\sin^{\beta+1}(x)}+\frac{\beta(\beta+1)x^{\beta}\cos^{2}(x)}{\sin^{\beta+2}(x)}$$

给出．在区间$\left[\frac{\pi}{2^{j+1}},\frac{\pi}{2^{j}}\right]$上，我们能用

$$\begin{aligned}u(x)&\approx u(m)-mu'(m)+\frac{m^{2}}{2}u''(m)+[u'(m)-mu''(m)]x+\frac{u''(m)}{2}x^{2}\\&\equiv a_{0}+a_{1}x+a_{2}x^{2}\end{aligned}$$

来逼近 $u(\cdot)$，其中 $m\equiv 3\pi/2^{j+2}$是区间中点．这产生

$$\begin{aligned}C_{j}\approx\tilde{C}_{j}&\equiv\frac{\sigma_{\varepsilon}^{2}2^{j+1}}{\pi 2^{\beta}}\int_{\pi/2^{j+1}}^{\pi/2^{j}}a_{0}x^{-\beta}+a_{1}x^{1-\beta}+a_{2}x^{2-\beta}\mathrm{d}x\\&=\frac{\sigma_{\varepsilon}^{2}2^{j+1}}{\pi 2^{\beta}}(A_{0,j}+A_{1,j}+A_{2,j}),\end{aligned}\tag{354}$$

其中

$$A_{n,j}\equiv a_{n}\int_{\pi/2^{j+1}}^{\pi/2^{j}}x^{n-\beta}\mathrm{d}x=\begin{cases}a_{n}\left(\dfrac{\pi}{2^{j+1}}\right)^{n+1-\beta}\dfrac{2^{n+1-\beta}-1}{n+1-\beta}, & \beta\neq n+1;\\ a_{n}\log(2), & \beta=n+1.\end{cases}$$

[5]对于分形的布朗运动(FBM)使用规范正交小波分解来近似变白滤波器在Flandrin(1992)已经讨论过．通过抽样分形布朗运动感兴趣的形式的离散序列，Flandrin 得出了离散计算．Tewfik and Kim (1992)在分析由抽样分形布朗运动的形式产生的小波系数的相关性时注意到了周期边界条件的影响，这些在他们的定理中无任何暗示．在某种条件下，得到边界系数不存在这种理想化假设，Dijkerman and Mazumdar (1994)证明了分形布朗运动的离散小波系数通过尺度按照指数快速衰减并且依时间按照双曲线快速衰减．分形布朗运动的更进一步的结果在 Masry(1993)和 Wang (1996)中可以看到，他们也讨论了分形布朗运动的"导数"(离散时间分形高斯噪声(FGN)过程的一个连续时间模拟)．

[6]用小波型变换来分析长记忆时间序列的思想至少可以追溯到 Allan (1966)，Allan 的"Allan 方差"用于高性能振荡器的频率平稳性的一个时域测量．正如 8.6 节所注意到的，当使用哈尔小波滤波器时，在特殊的尺度上 Allan 方差直接和小波系数在该尺度上的方差有关．对于平稳增量的过程可以计算出 Allan 方差．一个普通过程模型的使用是分形布朗运动，这是一个有平稳增量的连续时间过程(参见 7.6 的讨论)．这样的过程在一个宽的频率范围上有幂规律形式具有指数 α 的谱(Flandrin，1989)，并且指数和在特殊尺度上的 Allan 方差有简单的关系，因此提供了估计 α 的一种方法，正如 Percival(1983)所研究的．虽然仅当用哈尔小波滤波器时 Allan 方差和小波系数的方差有关，如果使用其他的适当紧支撑滤波器，那么在一个尺度上的小波系数的方差能再次和谱指数有关——参见 Abry 等(1993)的分形布朗运动例子，及 Percival(1995)．

9.2 长记忆过程的模拟

让我们现在考虑如何能使用离散小波变换近似去相关的分形差分过程以便产生该过程的模拟(本节和下一节的材料是建立在 Wornell(1993, 1996), McCoy and Walden(1996)之上的). 假设向量 $\mathbf{X}=[X_0, \cdots, X_{N-1}]^T$ 包含来自一个平稳分形差分过程$\{X_t\}$长为 $N=2^J$ 的一部分,该过程具有零均值和参数 δ 及 σ_ε^2. 令 $\mathcal{W}_N$ 是定义完全离散小波变换的 $N\times N$ 阶矩阵(为了使本节的讨论更清楚,对于这个矩阵把下标增加到这个矩阵的标准记号中). 如果我们令 $\Sigma_{\mathbf{X}}$ 是 $\mathbf{X}$ 的协方差矩阵,那么,如前面所述,对于离散小波变换 $\mathbf{W}=\mathcal{W}_N\mathbf{X}$ 的协方差矩阵由 $\Sigma_{\mathbf{W}}=\mathcal{W}_N\Sigma_{\mathbf{X}}\mathcal{W}_N^T$ 给出(这是式(262c)的应用). 去相关性质说明 $\Sigma_{\mathbf{W}}$ 的非对角元是相当小的. 正如 McCoy and Walden(1996)中,让我们用 $N\times N$ 阶对角阵 Λ_N 来近似这个矩阵,Λ_N 的对角元素为

$$\underbrace{C_1,\cdots,C_1}_{\frac{N}{2}\text{个}},\underbrace{C_2,\cdots,C_2}_{\frac{N}{4}\text{个}},\cdots,\underbrace{C_j,\cdots,C_j}_{\frac{N}{2^j}\text{个}},\cdots,\underbrace{C_{J-1},C_{J-1}}_{2\text{个}},C_J,C_{J+1}. \tag{355a}$$

正如式(343b),此处,$C_j\approx\operatorname{var}\{W_{j,t}\}$, $j=1,\cdots,J$, 而 C_{J+1} 仍然是由 $\mathbf{W}$ 的最后元素的方差决定的近似(即 $\mathbf{V}_J$ 的单个元素 $V_{J,0}$). 从练习[97]我们知道,对于任何基于 Daubechies 小波滤波器的离散小波变换,$\mathcal{W}_N$ 的最后一行的元素均等于 $1/\sqrt{N}$, 由此得出 $V_{J,0}$必须等于 $\overline{X}/\sqrt{N}$, 其中 $\overline{X}$ 是 $\mathbf{X}$ 的样本均值.

现在我们能通过下面的步骤获得高斯分形差分过程的一个近似模拟. 首先,可以用式(343c)通过数值积分计算近似方差 C_j, $j=1,\cdots,J$(另外可以用准确的方差 $\operatorname{var}\{W_{j,t}\}$, 它由式(343a)计算得到,来代替这些近似方差或者用式(344)给出的对 C_j 的简单近似$\tilde{C}_j$). 其次,利用下面的结果来决定 C_{J+1}(C_{J+1}是 $\operatorname{var}\{V_{J,0}\}$ 的近似).

练习[355] 证明

$$\operatorname{var}\{X_t\}=\frac{\operatorname{var}\{V_{J,0}\}}{N}+\sum_{j=1}^{J}\frac{\operatorname{var}\{W_{j,t}\}}{2^j}.$$

提示:从式(302)出发. ◁

如果把式(284c)给出的 $\operatorname{var}\{X_t\}$的表达式和对 $\operatorname{var}\{W_{j,t}\}$已经计算出的近似量 C_j 结合起来,就得到

$$C_{J+1}\equiv N\left(\frac{\sigma_\varepsilon^2\Gamma(1-2\delta)}{\Gamma^2(1-\delta)}-\sum_{j=1}^{J}\frac{C_j}{2^j}\right). \tag{355b}$$

最后,如果令 $\mathbf{Z}_N$ 表示偏离具有零均值和单位方差的高斯白噪声过程的 N 维向量,我们能形成

$$\mathbf{Y}_N\equiv\mathcal{W}_N^T\Lambda_N^{1/2}\mathbf{Z}_N,$$

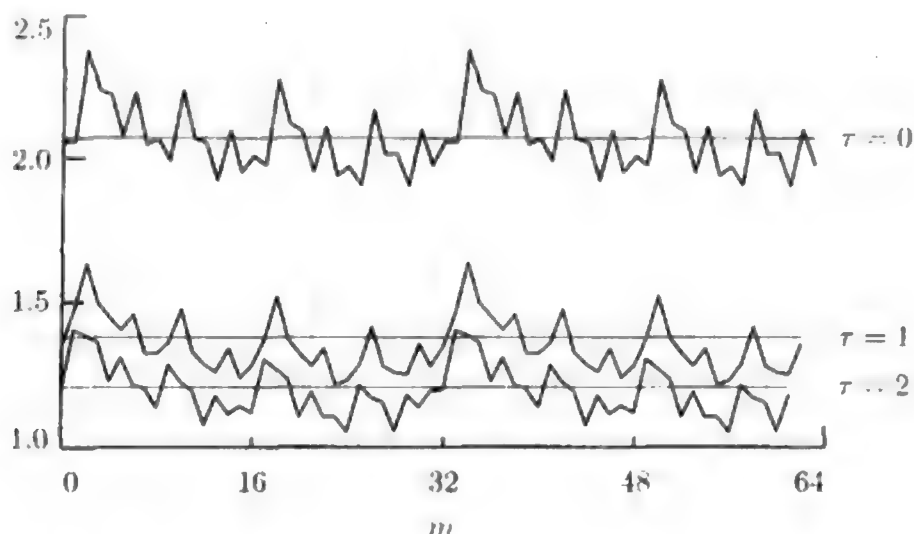


图 356 协方差矩阵 Σ_X 和 Σ_Y（分别是锯齿状粗线和水平细线）的对角元素 $\Sigma_{Y,m,m+\tau}$ 和 $\Sigma_{X,m,m+\tau}$，$m=0$，…，$N-1-\tau$，Σ_X 对于抽样尺寸 $N=64$ 由具有 $\sigma_\varepsilon^2=1$ 的一个 FD(0.4)过程而来，并且 Σ_Y 是使用 LA(8)离散小波变换构造的．对于每个协方差矩阵，绘出了三条对角线，即主对角线（$\tau=0$）和前两个偏对角线（$\tau=1$，2）．因此 Σ_X 显示了对平稳过程的 Toeplitz 结构，它没有近似 Σ_Y

我们把 $\mathbf{Y}_N$ 作为对 $\mathbf{X}$ 的一个近似（在上面，$\Lambda_N^{1/2}$ 是一个对角阵，其元素是 Λ_N 元素的平方根，即 $\Lambda_N^{1/2}\Lambda_N^{1/2}=\Lambda_N$）．

现在让我们来研究如何较好地把 $\mathbf{Y}_N$ 和 $\mathbf{X}$ 的统计属性匹配好．因为 $\mathbf{Y}_N$ 是高斯随机变量的一种线性组合，所以服从多变量高斯分布，正如我们假设的，对于 $\mathbf{X}$ 这也是真的．这个分布完全由它的均值向量和协方差矩阵决定．因为 $E\{\mathbf{Z}_N\}=\mathbf{0}$（$N$ 个零的向量），由此得出 $E\{\mathbf{Y}_N\}=\mathbf{0}$，这与分形差分过程 $\{X_t\}$ 的零均值假设是一致的．由于对 Z_N 的协方差矩阵正好是单位阵，式(262c)指出 $\mathbf{Y}_N$ 的协方差矩阵由 $\Sigma_Y=\mathcal{W}_N^T\Lambda_N\mathcal{W}_N$ 给出，这就是想尽可能近地逼近 Σ_X．现在 $\{X_t\}$ 是平稳过程，因此 Σ_X 的第(m，n)个元素由 $s_{X,m-n}$ 给出，其中 $\{s_{X,\tau}\}$ 为过程的自协方差序列．对于任何平稳过程，协方差矩阵具有 Toeplitz 结构，因为沿着任何给定的对角线上的值是相同的．计算指出 Σ_Y 没有这样的结构，即 Y_N 中的随机变量并不是平稳过程的一部分．作为例子（$\Sigma_{Y,m,n}$ 表示了 Σ_Y 的第(m，n)个元素），图 356 中锯齿状粗线显示了在 $N=64$，LA(8)离散小波变换和 $\delta=0.4$，$\sigma_\varepsilon^2=1$ 的分形差分过程的情况下 $\Sigma_{Y,m,m+\tau}$ 在 $m=0$，1，…，$N-1-\tau$($\tau=0$，1，2)时的图像．重叠的水平细线给出了 Σ_X 对应的值，即 $s_{X,0}$、$s_{X,1}$ 和 $s_{X,2}$．

虽然 $\mathbf{Y}_N$ 不是平稳的，但是应用随机循环平移，可以从 $\mathbf{Y}_N$ 产生一个平稳过程（如第 4 章所述，离散小波变换自然可以处理循环的时间序列）．相应地，令 κ 是一个在整数 0，1，…，$N-1$ 上均匀分布的随机变量，并且 $\widetilde{\mathbf{Y}}_N\equiv\mathcal{T}^{-\kappa}\mathbf{Y}_N$，其中 $\mathcal{T}$ 是围绕式(52a)的讨论定义的循环平移矩阵．练习[9.4]的责任是证明 $\widetilde{\mathbf{Y}}_N$ 中的

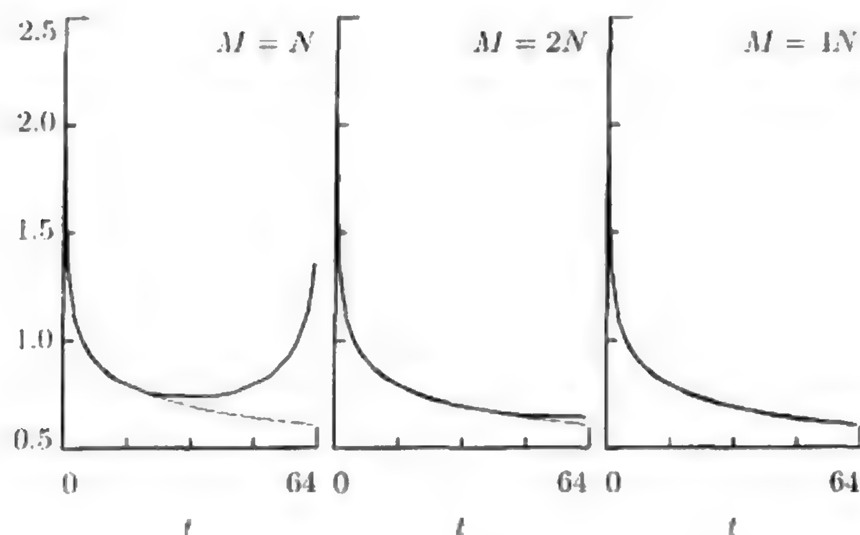


图 357 对于 FD(0.4)过程，真实的自协方差序列(细线)和基于小波的近似自协方差序列(粗线). 近似的自协方差序列是基于 LA(8)离散小波变换，在这个变换中，我们产生长为 M 的一个序列，之后扩展这个序列为长 $N=64$. 随着 M 从 N 变化到 $4N$，近似自协方差序列越来越接近于真实的自协方差序列

随机变量形成了平稳过程的一部分，该过程的自协方差序列由下式确定：

$$s_{\widetilde{Y},\tau} \equiv \frac{1}{N}\sum_{m=0}^{N-1} \Sigma_{Y,m,m+\tau \bmod N}, \quad \tau = 0,1,\cdots,N-1. \tag{357}$$

借助于图 356，通过平均化显示对角线可以得到自协方差序列(对于 $\tau=1$ 有一个附加项，即 $\Sigma_{Y,N-1,0}$，并且对于 $\tau=2$ 有两个附加项，即 $\Sigma_{Y,N-2,0}$ 和 $\Sigma_{Y,N-1,1}$). 图 357 的左图比较了目标 FD(0.4)过程(细线)和近似 $\widetilde{\mathbf{Y}}_N$ 过程(粗线)延迟 $\tau=63$ 的自协方差序列. 当对于小延迟有好的对应时，在中等和大的延迟上这并不正确(练习[9.4]的最后部分给出了这种差异性的解释：正如图形所表现的，事实上，对 $\widetilde{\mathbf{Y}}_N$ 的自协方差序列约束是关于 $\tau=N/2$ 对称，即对于 $\tau=1$，…，$N/2$ 有 $s_{\widetilde{Y},N-\tau}=s_{\widetilde{Y},\tau}$).

求助于基于离散傅里叶变换的高斯谱综合方法(GSSM，在 7.8 节叙述)的相同策略，我们可以获得更好的对 $\mathbf{X}$ 的近似模拟，也就是产生长为 $M>N$ 的模拟序列并且利用前 N 个值. 本质上，子采样破坏了离散小波变换所强加的内在循环性，因此导致了对要求的自协方差序列延迟 $\tau=N-1$ 的较好的近似. 图 357 的中间图和右图比较了对于长为 $N=64$ 的序列所要求的自协方差序列和近似的自协方差序列(分别是细线和粗线)，该序列分别是长为 $M=2N=128$ 和 $M=4N=256$ 的模拟序列的子采样. 在 $M=2N$ 情况下，在最大的延迟上仍然有某种微小差异，而 $M=4N$ 时有相当大的整体一致性. 对于 M 的选择和高斯谱综合方法中运行较好的一样，因此在这里也使用它.

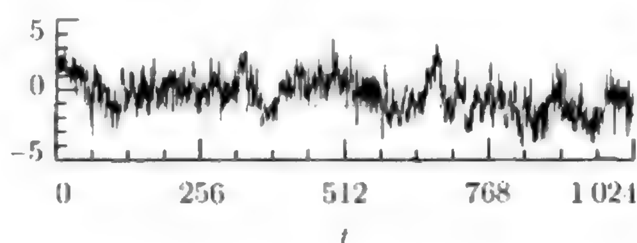


图 358 具有零均值和参数 $\delta=0.4$，$\sigma_\varepsilon^2=1.0$ 的 FD 过程采用基于 LA(8) 小波，长为 $N=1\,024$ 的序列的模拟

总之，用于产生长为 $N=2^J$ 的时间序列所使用的基于小波的过程建议包括如下几个步骤，而时间序列是分形差分过程的一个近似.

[1]在计算机上使用伪随机数发生器产生向量 $\mathbf{Z}_M$，该向量包含具有零均值和单位方差 $M=4N$ 的高斯白噪声偏离.

[2]使用数值积分，用式(343c)来计算近似方差 C_j，$j=1$，…，$J+2$. 在式两边用$J+2$代替 J，用式(355b)计算 C_{J+3}.

[3]用$\sqrt{C_1}$乘以 $\mathbf{Z}_M$ 中前 $M/2$ 个元素，用$\sqrt{C_2}$乘以接下来的 $M/4$ 个元素，依次类推，直到最后四个元素，最后四个元素分别用$\sqrt{C_{J+1}}$，$\sqrt{C_{J+1}}$，$\sqrt{C_{J+2}}$，$\sqrt{C_{J+3}}$来乘(当抽样尺寸由 N 增加到 $M=4N=2^{J+2}$时，这与式(355a)是一致的). 可以把乘法结果表示为 $\Lambda_M^{\frac{1}{2}}\mathbf{Z}_M$.

[4]计算 $\mathbf{Y}_M=\mathcal{W}_M^{\mathrm{T}}\Lambda_M^{\frac{1}{2}}\mathbf{Z}_M$(实际上，我们不明显地生成$\mathcal{W}_M$，而是使用第 4.6 节评论与扩展中描述的离散小波逆变换塔式算法).

[5]利用伪随机数发生器得到在整数 0，1，…，$M-1$ 上的从一致分布的一个偏离 κ. 然后想要的模拟序列是 Y_κ，$Y_{\kappa+1 \bmod M}$，…，$Y_{\kappa+N-1 \bmod M}$，其中 Y_t 是 $\mathbf{Y}_M$ 的第 t 个元素.

例如，图 358 所示的是长为 $N=1\,024$ 的一个序列，这个序列是从使用上面带有 LA(8)小波的程序的一个 FD(0.4)过程模拟得到的. 定性地，这个序列与图 342 的底部的一个序列在性质上是类似的，它已经用在 7.8 节中描述的基于离散傅里叶变换的精确的 Davies-Harte 方法生成的.

实际检验证明了在正确的性质下基于小波的方法在仿真中如何得好. 为了研究这些，我们用 LA(8)基于小波的方法和 Davies-Harte 方法生成的一个长为$N=64$ 的 10 000 个模拟FD(0.4)序列. $\widetilde{X}_{k,t}$，$t=0$，1，…，$N-1$ 表示任何一个模拟方法的第 k 个序列. 对于每个序列，我们计算自协方差序列的一个无偏估计：

$$\hat{s}_{k,\tau}^{(u)}\equiv\frac{1}{N-\tau}\sum_{t=0}^{N-1-\tau}\widetilde{X}_{k,t}\widetilde{X}_{k,t+\tau},\quad \tau=0,\cdots,N-1,$$

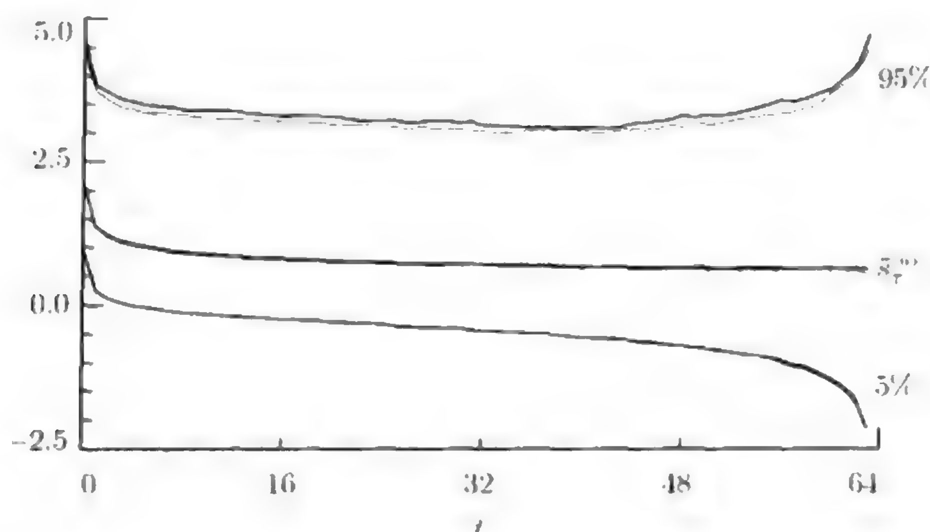


图 359　Davies-Harte 方法(图中部细线)和基于 LA(8)小波方法(粗线)对于 FD(0.4)过程生成的超过 10 000 个实现的平均的自协方差矩阵的估计．图中上下两条曲线表示用两种方法时$\hat{s}_{\tau}^{(u)}$ 的 5%和 95%的 10 000 个模拟的经验分布

然后形成$\hat{s}_{k,\tau}^{(u)}$在所有 10 000 实现上的抽样均值：

$$\bar{s}_{\tau}^{(u)} \equiv \frac{1}{10\,000}\sum_{k=0}^{9\,999}\hat{s}_{k,\tau}^{(u)}.$$

两种方法得出的序列 $\bar{s}_{\tau}^{(u)}$ 如图 359 的中间部分所示，其中粗线和细线分别表示基于小波的方法和 Davies-Harte 方法得出的结果．这两条曲线实质上是一样的(这个图所给出的分辨，真实的自协方差序列的图的一部分和 Davies-Harte 方法$\bar{s}_{\tau}^{(u)}$是不能区分开来的)．图中上下两条曲线表示用两种方法时$\hat{s}_{k,\tau}^{(u)}$的 5%和 95%的经验分布．在用基于小波的方法时低的百分点很接近，在高的百分点上有明显的一点向上的偏离，基于小波的方法的演示完全可以接受．

因为基于离散傅里叶变换的 Davies-Harte 方法可以得出具有正确的统计性质的高斯分形差分过程的模拟，它比任何近似的方法都优越，甚至和基于小波的方法一样精确，然而基于小波的方法可以容易地适合提供在两种情况下的模拟，而用 Davies-Harte 方法的时候是会出现问题的：模拟很长的时间序列(例如，在某些实时的应用中是需要的)和模拟具有随时间展开随机性质的序列．

在第一种情况下，用 Davies-Harte 方法来生成一个很长的时间序列是有问题的，因为必须用典型的快速傅里叶变换算法来计算离散傅里叶变换 $N=10^7$ 个值，这是计算机时间和记忆禁止的量．快速傅里叶变换算法和离散小波变换塔式算法是 $O(N\log_2(N))$和 $O(N)$的计算复合，因此，当 N 很大的时候离散小波变换算法具有很大的时间优势．另外，如果一个很长的序列对实时的应用来说是必要的，那么离散小波逆变换算法的局部性质可以开发成有效地模拟时间序列"起飞"．为了给出此项工作的一个想法，我们考虑，比如 $\mathbf{Y}_M=\mathcal{W}_M^{\mathrm{T}}\mathbf{W}$ 的元素 Y_{97}，

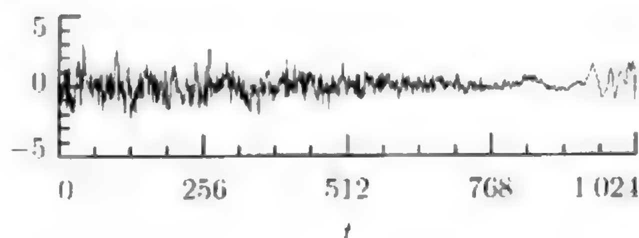


图 360 基于 LA(8)小波方法模拟由过程的一列长度 $N=1\,024$ 的序列，该过程具有时间变化统计的性质

其中 $\boldsymbol{W}\equiv\Lambda_M^{\frac{1}{2}}\boldsymbol{Z}_M$，且我们假设 $M=2^{J+2}\gg97$．因为依照 $\boldsymbol{Y}_M$，$\boldsymbol{W}$ 包含离散小波变换系数，所以我们可以把它分成通常的分量 $\boldsymbol{W}_1$，…，$\boldsymbol{W}_{J+2}$ 和 $\boldsymbol{V}_{J+2}$．如果我们专门研究哈尔离散小波变换，就可以用 $\boldsymbol{W}$ 中的 $J+3$ 个元素的线性组合来表示 Y_{97}，即子向量的一种形式可以是如下方式：

$$W_{1,48},W_{2,24},W_{3,12},W_{4,6},W_{5,3},W_{6,1},W_{7,0},\cdots,W_{J+2,0},V_{J+2,0}.$$

当我们从 Y_{97} 运动到 Y_{98} 时，线性组合解出几乎相同的一组元素，我们只需用 $W_{1,49}$ 代替 $W_{1,48}$．另外 Y_{97} 和 Y_{98} 的前两个系数保持相等$\left(\text{分别由}\frac{1}{\sqrt{2}}\text{和}-\frac{1}{4}\text{变成}-\frac{1}{\sqrt{2}}\text{和}\frac{1}{4}\right)$．如所描述的，因此我们能够借助于 Y_{97} 来表示 Y_{98}：

$$Y_{98}=Y_{97}-\frac{W_{1,48}+W_{1,49}}{\sqrt{2}}+\frac{W_{2,24}}{2}.$$

当我们继续下去，在每个时间点能够继续直接用每个子向量中的一个元素．记忆需要是十分节制的，并且这个方法对于满足实时的计算的需要已经很简单了．相关的方法能够提出，例如D(4)和 LA(8)小波，由于 Y_t 依赖于每个子向量中的少数元素(Percival and Constantine，2000)．

现在让我们简单地讨论一下，基于小波的模拟方法如何能够适应模拟其统计性质变化超时的时间序列．在模拟很长的时间序列的情况下，主要的是在 $\boldsymbol{Y}_M$ 中已给定的元素 $\boldsymbol{W}$ 中的每个元素有部分的影响．作为一个简单的例子，我们假设 $M=N=2^J$，在每个式(355a)的 $J+1$ 组中生成对角矩阵 Λ_N 的元素，因此 C_t 值由应用于一个 FD(0.4)过程的 LA(8)离散小波变换来决定，但是我们可以修正对组中第 t 个元素，由$(N_j-t)C_j/N_j$($t=0$，…，N_j-1)给出，其中 $j=1$，…，4．这模仿了这样一个过程，这个过程的高尺度的性质用长记忆过程来决定，但是它的低尺度的性质随着时间而变化．这个过程的一种实现已经在图 360 中展示了(为了精确地控制一个过程的统计性质合拍的演变，我们需要知道哪个小波参数决定时间序列的哪一部分．时间与 $\boldsymbol{W}_j$ 中每个参数之间的关系对于哈尔小波很容易做出，这在 4.8 节中已经发展到对于最接近对称小波的情形了，为了保持这个例子的简单性，我们就不使用这些关系了)．

最后，我们注意 Davies-Harte 需要自协方差序列的知识，它是真正有利于分形差分过程，用来测定另外的过程并不容易，基于小波的模拟工作完全需要平均谱密度函数的层或者等价的小波方差：没有必要知道过程的协方差矩阵．因此我们可以思考，对于标准的短记忆过程可以将这些表达为一个普通的模拟技术处理，从而一个完全已知的协方差矩阵适于 Cholesky 分解（例如，参见 McLeod and Hipel(1978)或者 Kay，1981)．

9.2 节的评论与扩展

[1]用小波按照不相关的具有一个已知方差级数的小波系数来解释合成的随机过程的思想是由 Wornell(1990，1993)提出的．这些文章集中于模拟连续的非平稳的过程，这些过程当 $f\to 0$ 时具有幂规律性质．实际上，模拟方案再次规划为一个迭代"上抽样滤波器融合"的离散时间序列算法，后面是单一的离散-连续卷积(Wornell，1993，附录 I)．Masry(1993)将 Wornell 方法应用于分形布朗运动时对他的方法进行了评论：他指出分形布朗运动不能够经过一个小波级数表示严密地合成，因为对于分形布朗运动小波系数不是不相关的（然而，看起来 Wornell 总是打算使他的方法变成一个近似的结果)．

[2]有趣的是，我们注意到$\widetilde{\mathbf{Y}}_N$ 的任意元素$\widetilde{Y}_t$ 的单变量分布是高斯混合模型，这一概念我们将在 10.4 节中介绍和利用，随机的循环平移多变量高斯随机变量的一个向量的方法，因此是构造具有非高斯的单变量分布的平稳过程的一种可行的方法．

9.3　平稳分形差分过程的极大似然估计

假设现在我们有一个时间序列是具有未知参数 $\delta\in\left[-\frac{1}{2},\ \frac{1}{2}\right)$ 和 $\sigma_\varepsilon^2>0$ 的零均值的平稳分形差分过程，是由 $\mathbf{X}=[X_0,\ \cdots,\ X_{N-1}]^T$ 部分实现，其中如前 $N\equiv 2^J$ 对于某个整数 J 成立（即使最初我们最感兴趣的是长记忆分形差分过程，即 $\delta\in\left(0,\ \frac{1}{2}\right)$，这里出现的理论在 $\delta\in\left[-\frac{1}{2},\ 0\right]$的短记忆分形差分过程中也成立)．我们想基于 $\mathbf{X}$ 来估计这些参数．在假设 $\mathbf{X}$ 服从多变量的高斯（正态)分布的情况下，使用极大似然方法来估计 δ 和 σ_ε^2．给定观测值 $\mathbf{X}$，对于 δ 和 σ_ε^2 的似然函数由

$$L(\delta,\sigma_\varepsilon^2\mid\mathbf{X})\equiv\frac{1}{(2\pi)^{N/2}\mid\Sigma_{\mathbf{X}}\mid^{1/2}}\mathrm{e}^{-\mathbf{X}^T\Sigma_{\mathbf{X}}^{-1}\mathbf{X}/2}\qquad(361)$$

给出，其中 $\Sigma_{\mathbf{X}}$ 是 $\mathbf{X}$ 的协方差矩阵（即它的第$(m,\ n)$个元素是由式(284b)的 $s_{X,m-n}$给出的)，而 $\mid\Sigma_{\mathbf{X}}\mid$ 表示 $\Sigma_{\mathbf{X}}$ 的行列式．注意似然函数是独自通过 $\Sigma_{\mathbf{X}}$ 依赖于 δ 和 σ_ε^2的．我们能够对于任何特殊参数值进行主要的关于似然函数的估计．δ 和 σ_ε^2 的极大似然估计是 $L(\delta,\ \sigma_\varepsilon^2\mid\mathbf{X})$作为 δ 和 σ_ε^2 函数的最大值．但是，存在着两个准确的极大似然估计的实际问题(Beran，1994，5.4 节)．首先它们的行列式的计算是很耗时的，因为 $L(\delta,\ \sigma_\varepsilon^2\mid\mathbf{X})$是由估计来计算的，这是有损耗的，甚至

对于中等的 N 也是如此. 其次，当 δ 趋于 $\frac{1}{2}$ 时在计算似然方程时有可能是数值不平稳的. 有效的数值稳定的似然估计是我们考虑且感兴趣的.

这里我们考虑一个近似的极大似然估计的方法. 实际上 Σ_X 能够由 $\widetilde{\Sigma}_X \equiv \mathcal{W}^T \Lambda_N \mathcal{W}$ 近似给出，其中 $\mathcal{W}$ 是一个离散小波变换矩阵，Λ_N 是一个以 C_j 为对角元素的对角阵，其中 C_j 在式(355a)中描述. 由这个近似，我们有

$$L(\delta,\sigma_\varepsilon^2 \mid \boldsymbol{X}) \approx \widetilde{L}(\delta,\sigma_\varepsilon^2 \mid \boldsymbol{X}) \equiv \frac{1}{(2\pi)^{N/2} \mid \widetilde{\Sigma}_X \mid^{1/2}} \mathrm{e}^{-\boldsymbol{X}^T \widetilde{\Sigma}_X^{-1} \boldsymbol{X}/2}. \tag{362a}$$

现在 $\widetilde{L}(\delta, \sigma_\varepsilon^2 \mid \boldsymbol{X})$ 的最大值等价于(重新尺度化与重新局部化)对数似然函数

$$\widetilde{l}(\delta,\sigma_\varepsilon^2 \mid \boldsymbol{X}) \equiv -2\log(\widetilde{L}(\delta,\sigma_\varepsilon^2 \mid \boldsymbol{X})) - N\log(2\pi) = \log(\mid \widetilde{\Sigma}_X \mid) + \boldsymbol{X}^T \widetilde{\Sigma}_X^{-1} \boldsymbol{X} \tag{362b}$$

的最小值. 我们可以用下面的结果简化第一项.

练习[362a] 如上 $N_j \equiv N/2^j$，证明

$$\log(\mid \widetilde{\Sigma}_X \mid) = \log(C_{J+1}) + \sum_{j=1}^{J} N_j \log(C_j). \quad \triangleleft$$

在这点上要了解 C_j 是怎么精确地依赖于 δ 和 σ_ε^2 是很方便的. 对 $j=1, \cdots, J$，式(343c)表示，我们能写成

$$C_j = \sigma_\varepsilon^2 C'_j(\delta) \quad \text{其中} \quad C'_j(\delta) \equiv 2^{j-1} \int_{1/2^{j+1}}^{1/2^j} \frac{1}{[4\sin^2(\pi f)]^\delta} \mathrm{d}f; \tag{362c}$$

另一方面 $j=J+1$，由式(355b)可得

$$C_{J+1} = \sigma_\varepsilon^2 C'_{J+1}(\delta), \quad \text{其中} \quad C'_{J+1}(\delta) \equiv N \frac{\Gamma(1-2\delta)}{\Gamma^2(1-\delta)} - \sum_{j=1}^{J} N_j C'_j(\delta). \tag{362d}$$

现在我们可以写成

$$\log(\mid \widetilde{\Sigma}_X \mid) = N\log(\sigma_\varepsilon^2) + \log(C'_{J+1}(\delta)) + \sum_{j=1}^{J} N_j \log(C'_j(\delta)).$$

式(362b)第二项可以如下化简.

练习[362b] 证明

$$\boldsymbol{X}^T \widetilde{\Sigma}_X^{-1} \boldsymbol{X} = \frac{1}{\sigma_\varepsilon^2}\left(\frac{V_{J,0}^2}{C'_{J+1}(\delta)} + \sum_{j=1}^{J} \frac{1}{C'_j(\delta)} \sum_{t=0}^{N_j-1} W_{j,t}^2\right). \quad \triangleleft$$

表 363

δ		MLE			
		哈尔	D(4)	LA(8)	确切值
0.25	均值	0.218 4	0.229 3	0.232 8	0.237 4
	偏差	−0.031 6	−0.020 7	−0.017 2	−0.012 6
	SD	0.071 3	0.070 5	0.071 0	0.067 3
	RMSE	0.078 0	0.073 5	0.073 1	0.068 5
0.4	均值	0.361 4	0.372 7	0.376 8	0.379 7
	偏差	−0.038 6	−0.027 3	−0.023 2	−0.020 3
	SD	0.067 5	0.065 2	0.064 0	0.060 4
	RMSE	0.077 8	0.070 7	0.068 1	0.063 7

注：基于利用哈尔、D(4)和 LA(8)小波滤波器的式(362a)的似然函数，参数 δ 的基于小波的近似极大似然估计$\tilde{\delta}^{(s)}$的抽样均值、偏差、标准差(SD)和平方误差均值的平方根(RMSE). 所有 1 024 个时间序列都取长度 $N=128$ 和模拟使用 Davies-Harte 方法. 相应的确切极大似然估计$\hat{\delta}$统计量在最后一列给出.

对 $\log(|\tilde{\Sigma}_X|)$和 $\boldsymbol{X}^T\tilde{\Sigma}_X^{-1}\boldsymbol{X}$ 的表达形式如上，我们可以把式(362b)重新表示为

$$\tilde{l}(\delta,\sigma_\varepsilon^2\mid\boldsymbol{X})=N\log(\sigma_\varepsilon^2)+\log(C'_{J+1}(\delta))+\sum_{j=1}^{J}N_j\log(C'_j(\delta))$$
$$+\frac{1}{\sigma_\varepsilon^2}\left(\frac{V_{J,0}^2}{C'_{J+1}(\delta)}+\sum_{j=1}^{J}\frac{1}{C'_j(\delta)}\sum_{t=0}^{N_j-1}W_{j,t}^2\right).\tag{363a}$$

对式(363a)的右边关于 σ_ε^2求微分，再使得到的表达式为零，我们可以得到一个 σ_ε^2的近似极大似然估计$\tilde{\sigma}_\varepsilon^2$ 的表达式. 关于 σ_ε^2求解得到估计值$\tilde{\sigma}_\varepsilon^2$，它能够作为 δ 的函数用

$$\tilde{\sigma}_\varepsilon^2(\delta)\equiv\frac{1}{N}\left(\frac{V_{J,0}^2}{C'_{J+1}(\delta)}+\sum_{j=1}^{J}\frac{1}{C'_j(\delta)}\sum_{t=0}^{N_j-1}W_{j,t}^2\right)\tag{363b}$$

给出. 利用上面的表达式，我们可以从式(363a)估计参数 σ_ε^2，得到 Brockwell and Davis(1991)提到的简化的对数似然值：

$$\tilde{l}(\delta\mid\boldsymbol{X})\equiv\tilde{l}(\delta,\tilde{\sigma}_\varepsilon^2(\delta)\mid\boldsymbol{X})-N$$
$$=N\log(\tilde{\sigma}_\varepsilon^2(\delta))+\log(C'_{J+1}(\delta))+\sum_{j=1}^{J}N_j\log(C'_j(\delta)).\tag{363c}$$

一旦我们把式(363b)带入上式，化简的似然值就仅仅依赖于单个参数 δ，而在区间 $\left[-\frac{1}{2},\frac{1}{2}\right)$上确定 δ 的数值，即$\tilde{\delta}^{(s)}$，是一个简单的问题. 最小化 $\tilde{l}(\delta\mid\boldsymbol{X})$得到$\tilde{\delta}^{(s)}$，这个最小值就是 δ 的近似的极大似然估计，在我们决定了$\tilde{\delta}^{(s)}$的值后，通过把它代入式(363b)，就可以得到相关的 σ_ε^2的近似极大似然估计$\tilde{\sigma}_\varepsilon^2$.

表 364

δ		MLE			
		哈尔	D(4)	LA(8)	确切值
0.25	均值	0.225 6	0.236 3	0.240 2	0.244 3
	偏差	−0.024 4	−0.013 7	−0.009 8	−0.005 7
	SD	0.050 5	0.049 5	0.050 2	0.047 9
	RMSE	0.056 1	0.051 4	0.051 1	0.048 3
0.4	均值	0.371 0	0.383 2	0.388 6	0.390 0
	偏差	−0.029 0	−0.016 8	−0.011 4	−0.010 0
	SD	0.048 8	0.047 8	0.046 5	0.043 7
	RMSE	0.056 7	0.050 6	0.047 9	0.044 8

注：如同表 363，但是现在取 $N=256$.

为了看看这些近似的极大似然估计效果如何，我们用 Davies-Harte 方法(参见 7.8 节)模拟 1 024 个高斯 FD(0.25)和 FD(0.4)长度为 $N=128$ 和 $N=256$ 的时间序列，然后使式(363c)极小来估计 δ 的值$\tilde{\delta}^{(s)}$，我们考虑 3 种不同的小波滤波器：哈尔、D(4)和 LA(8). 对于给定的 δ 和 N，令$\tilde{\delta}^{(s),n}$是第 n 个模拟序列的极大似然估计. 对于 δ 的两个估计值和 $N=128$，表 363 给出了$\tilde{\delta}^{(s),n}$的样本均值$\bar{\delta}^{(s)}$和样本偏差(即$\bar{\delta}^{(s)}-\delta$)，以及样本标准离差(SD)和平方误差均值的平方根(RMSE)，即

$$\left(\frac{1}{1\,024}\sum_{n=0}^{1\,023}(\tilde{\delta}^{(s),n}-\bar{\delta}^{(s)})^2\right)^{1/2} \quad 和 \quad \left(\frac{1}{1\,024}\sum_{n=0}^{1\,023}(\tilde{\delta}^{(s),n}-\delta)^2\right)^{1/2}.$$

表 364 显示了当 $N=256$ 时的相关结果. 对于 δ 的两种选择，哈尔小波导致比其他两种小波更大的偏差和平方误差均值的平方根. 当 $\delta=0.25$ 时，在 D(4)和 LA(8)的结果中实质上不存在选择：前者存在较大的偏差，但平方误差均值的平方根则几乎完全一致. 当 $\delta=0.4$ 时，LA(8)的确比 D(4)要更好一些. 我们可以想到，当我们把样本容量由 128 增至 256 时，偏差和平方误差均值的平方根都会下降.

两个表格的最后一列都给出了 δ 的精确极大似然估计$\hat{\delta}$的相关结果(参见评论与扩展第[4]条)，我们看到 LA(8)小波的平方误差均值的平方根比精确方法大约大 10%，所以基于小波的逼近是相当好的. Kashyap and Eom(1988)导出了对于 δ 无偏估计的均方误差的渐近，Gramér-Rao 下界$\frac{6}{N\pi^2}$，当 $N=128$ 和 256 时，平方误差均值的平方根分别为0.068 9和 0.048 7. 这些值对 D(4)、LA(8)和精确方法而言都是与平方误差均值的平方根相符合的.

我们的结论是到目前为止都假设 $\boldsymbol{X}$ 是零均值平稳分形差分过程实现的一部分. 在实际应用中，我们很少知道过程对于一个时间序列先验的均值 $\mu\equiv E\{X_t\}$，所以有趣的是，修改我们的计划使 μ 成为另外一个待估计的参数. 因此，给定观测值 $\boldsymbol{X}$，我们现在考虑将似然作为 δ、σ_ε^2 和 μ 的函数，使式(361)变成

$$L(\delta,\sigma_\varepsilon^2,\mu\mid\boldsymbol{X})\equiv\frac{1}{(2\pi)^{N/2}\mid\Sigma_{\boldsymbol{X}}\mid^{1/2}}\mathrm{e}^{-(\boldsymbol{X}-\mu\boldsymbol{1})^{\mathrm{T}}\Sigma_{\boldsymbol{X}}^{-1}(\boldsymbol{X}-\mu\boldsymbol{1})/2},$$

其中 $\mathbf{1}$ 是 N 维向量，它的所有元素都是 1. 式(362b)的对数似然函数现在变成

$$\tilde{l}(\delta,\sigma_\varepsilon^2,\mu \mid \boldsymbol{X}) = \log(|\tilde{\Sigma}_{\boldsymbol{X}}|) + (\boldsymbol{X}-\mu\mathbf{1})^{\mathrm{T}}\tilde{\Sigma}_{\boldsymbol{X}}^{-1}(\boldsymbol{X}-\mu\mathbf{1}).$$

第一项可以再次化简作为练习[362a]，而第二项则可以变成如下(参见练习[362b]).

练习[365a] 证明

$$(\boldsymbol{X}-\mu\mathbf{1})^{\mathrm{T}}\tilde{\Sigma}_{\boldsymbol{X}}^{-1}(\boldsymbol{X}-\mu\mathbf{1}) = \frac{1}{\sigma_\varepsilon^2}\left(\frac{N^2(\bar{X}-\mu)^2}{C'_{J+1}(\delta)} + \sum_{j=1}^{J}\frac{1}{C'_j(\delta)}\sum_{t=0}^{N_j-1}W_{j,t}^2\right). \quad \triangleleft$$

式(363a)现在变为

$$\tilde{l}(\delta,\sigma_\varepsilon^2,\mu \mid \boldsymbol{X}) = N\log(\sigma_\varepsilon^2) + \log(C'_{J+1}(\delta)) + \sum_{j=1}^{J}N_j\log(C'_j(\delta))$$

$$+\frac{1}{\sigma_\varepsilon^2}\left(\frac{N^2(\bar{X}-\mu)^2}{C'_{J+1}(\delta)} + \sum_{j=1}^{J}\frac{1}{C'_j(\delta)}\sum_{t=0}^{N_j-1}W_{j,t}^2\right),$$

从式中明显看到，μ 使上式极小化的值产生极大似然估计 $\tilde{\mu}\equiv\bar{X}$，对于 δ 与 σ_ε^2 的极小化值同样无关紧要. 因此 μ 一经确定，对于 σ_ε^2 的极大似然估计现在取

$$\tilde{\sigma}_\varepsilon^2(\delta) = \frac{1}{N}\sum_{j=1}^{J}\frac{1}{C'_j(\delta)}\sum_{t=0}^{N_j-1}W_{j,t}^2 \tag{365}$$

(参见式(363b)). 有了对 $\tilde{\sigma}_\varepsilon^2(\delta)$ 的新定义，对于 δ 的递减对数似然函数的形式和前面的相同(式(363c)).

练习[365b] 假设有用于时间序列 $\boldsymbol{X}$ 的计算程序并且在过程均值已知为 0 的假设下计算极大似然估计 $\tilde{\delta}^{(s)}$ 和 $\tilde{\sigma}_\varepsilon^2$. 讨论：如果 $\boldsymbol{X}$ 有一个未知过程均值 μ，则用程序对 $\boldsymbol{X}-\bar{X}\mathbf{1}$ 代替 $\boldsymbol{X}$ 能够获得对 δ 和 σ_ε^2 的极大似然估计(即仅需要在引用程序以前减去每次观察的抽样均值). ◁

为了看未知过程均值的效果，从新计算可以得到表 364，每个模拟 $\boldsymbol{X}$ 用 $\boldsymbol{X}-\bar{X}\mathbf{1}$ 代替，在表中有单一的差异. 表 366 给出了在 1 024 个模拟序列上，对 δ 的各种极大似然估计的统计总结. 偏差、标准离差和平方误差均值的平方根在这里都明显比表 364 中相应的序列差. 因此在试着去估计 δ 时，如果对过程均值缺乏了解，就要付出代价(注意在 8.1 节讨论长记忆过程的样本方差所得到的类似结果).

作为例子，使用具有未知均值的一个实际时间序列，参见 9.8 节，那里对尼罗河最低水位计算了 $\tilde{\delta}^{(s)}$.

表　366

δ		MLE			
		哈尔	D(4)	LA(8)	确切值
0.25	均值	0.205 8	0.218 2	0.222 7	0.227 4
	偏差	−0.044 2	−0.031 8	−0.027 3	−0.022 6
	SD	0.055 9	0.055 1	0.055 7	0.052 8
	RMSE	0.071 2	0.063 6	0.062 0	0.057 5
0.4	均值	0.344 9	0.360 2	0.367 2	0.368 7
	偏差	−0.055 1	−0.039 8	−0.032 8	−0.031 3
	SD	0.055 0	0.053 8	0.052 5	0.049 4
	RMSE	0.077 8	0.066 9	0.061 9	0.058 5

注：如同表 364，不同的是过程均值假设未知，因此使用样本均值 $\overline{X}$ 进行估计.

9.3 节的评论与扩展

[1]Wornell and Oppenheim(1992)、Wornell(1993)和 Wornell(1996)考虑了连续时间高斯 $1/f$ 型过程的 δ 和 σ_ε^2 的等价的估计. 他们借助于小波系数单纯地形成似然函数. 此处的方法包括尺度系数 $V_{J,0}$，这在$\{X_t\}$已知有零均值的假设下可以得到. 当此事为真时，系数 $V_{J,0}$ 包含了关于最低频带$[-1/2^{J+1}, 1/2^{J+1}]$中的带通方差的信息，没有这些则估计方案就是比较差的(参见表 364 和 366). 当$\{X_t\}$不知道有零均值时(在实际应用中是很常见的情况)，系数 $V_{J,0}$ 是带通方差的较差的偏差估计值，这就说明了为什么这个系数在 μ 由 $\overline{X}$ 估计出来之后形成 δ 的缩减对数似然函数时是跳跃的. 然而，注意到这个函数仍然依赖于贯穿项 $\log(C'_{J+1}(\delta))$ 的 $\text{var}\{V_{J,0}\}=C_{J+1}=\sigma_\varepsilon^2 C'_{J+1}(\delta)$ 这一情况是很有趣的. 这种依赖直观上是合理的，因为 C_{J+1} 在构成对 Σ_X 的基于小波的近似量$\widetilde{\Sigma}_X$ 时是一个重要成分(这里是尺度系数方差重要性的一种衡量：如果我们在构造图 357 的 $M=4N$ 基于小波近似量自协方差序列时用零来代替它，则得到的自协方差序列在延迟 0 的 1.41 下降到延迟 63 的−0.05，然而真实的自协方差序列是从 2.07 到 0.61).

Kaplan and Kuo(1993)讨论了 Wornell and Oppenheim(1992)方法的一种修改. 他们指出在较后的参考文献中：离散小波变换实际上用于离散分形布朗运动(即离散抽样连续时间分形布朗运动过程)，以至于含有尺度的依赖方差级数是有偏差的，导致了谱指数估计的潜在问题. 他们证明分形高斯噪声(离散分形布朗运动的一阶差分)的使用在仅用哈尔小波时导致了无偏差方差级数.

[2]Jensen(1999b, 2000)把基于小波的极大似然估计延伸到处理自回归的、分形积分的、移动平均的过程(ARFIMA). 对于分形差分过程的这些延拓包含 p 个自回归参数和 q 个移动平均参数，这是为了模拟时间序列的小尺度(高频)属性. 分形差分过程的具体评述是它们仅依赖于参数 δ 和 σ_ε^2. 本质上，δ 控制了当 $f\to 0$ 时，对谱密度函数的近似幂规律，而σ_ε^2 仅是谱水平，不影响谱密度函数的形状. 使用 δ 和 σ_ε^2 来捕捉大尺度(低频)特征，并且谱水平使我们没有方法去处理

小尺度性质——这些输入的方法不需要对特殊时间序列有好的匹配. 因此如果我们试着由时间序列估计在小尺度不违背分形差分模型的 δ 和 σ_ϵ^2，那么 δ 的极大似然估计有明显的偏差，这是因为最大似然方案要求在每个倍频程带中的谱水平要跟随着一个分形差分模型. 在一个自回归、分形积分、移动平均模型中的额外 $p+q$ 个参数允许小的和大的尺度性质被消除，并且 δ 相应的估计值确切地反映了一个时间序列的大尺度性质.（应该注意到，接下来的两节所讨论的 δ 的估计值容易调整来减少对不和谐小尺度的依赖性——参见 9.7 节的例子.）

分形差分过程的上述评论也适用于分形高斯噪声和完全幂规律过程，两者都正好依赖于两个参数(图 282 显示对分形高斯噪声、完全幂规律过程和分形差分过程参数对的恰当选择导致了频密度函数有相似低频覆盖和高度，但是有高频特征的是不可控制的，因此是任意的). 拘于这个局限性，Kaplan and Kuo(1994)提出了一个对分形高斯噪声的扩展，分形高斯噪声允许在小尺度有不同的相关性效果.

[3]为了比较下两节所提供的 δ 的估计值. 这里讨论基于小波的极大似然估计局限在样本尺寸 N 为 2 的幂的时间序列. 如果 N 不是这样的，围绕这种局限性的一个简单方案是平均 δ 的两种估计，δ 的计算是对于两个子序列 X_0，…，$X_{N'-1}$ 和 $X_{N-N'}$，…，X_{N-1}，其中 N' 是满足 $N'<N$ 的 2 的最大幂.

[4]作为结束，这里注意需要找到对有零均值的平稳高斯分形差分过程的确切极大似然估计(详情参见 Beran，1994，5.3 和 2.5 节，或者 Brockwell and Davis，1991，13.2 节). 给定特殊的 δ，首先计算部分自相关序列(PACS)$\phi_{t,t}$，$t=1$，…，$N-1$. 对于分形差分过程给定形式为 $\phi_{t,t}=\frac{\delta}{t-\delta}$(Hosking，1981). 我们说明这个部分自相关序列的两种用处. 第一，我们用它来递归计算 X_t 的最好线性预估计的剩余系数，对于 $t=2$，…，$N-1$，通过

$$\phi_{t,k}=\phi_{t-1,k}-\phi_{t,t}\phi_{t-1,t-k},\quad k=1,\cdots,t-1$$

(这些是关键的递推式，是著名的 Levinson-Durbin 递推)给出 X_{t-1}，…，X_0，用这些系数形成观察预测误差：

$$e_t\equiv X_t-\sum_{k=1}^{t}\phi_{t,k}X_{t-k},\quad t=1,\cdots,N-1$$

(为了方便，令 $e_0\equiv X_0$). 部分自相关序列的第二个使用是计算与方差 e_t 和 e_0 有关的序列$\{v_t\}$：

$$v_t=\operatorname{var}\{e_0\}\prod_{n=1}^{t}(1-\phi_{n,n}^2),\quad t=0,\cdots,N-1,$$

其中

$$\operatorname{var}\{e_0\}=\frac{\Gamma(1-2\delta)}{\Gamma^2(1-\delta)}$$

(参见式(284c))；这里 $v_0=\operatorname{var}\{e_0\}=\operatorname{var}\{X_t\}$. 给定 $\mathbf{X}$，序列$\{\phi_{t,k}\}$、$\{e_t\}$和$\{v_t\}$都是 δ 的隐函数并且完全由它决定.

我们用来计算精确极大似然估计的公式和那些用来计算基于小波的近似极大似然估计的公式几乎相等. 代替式(363b)中近似最大似然估计值$\tilde{\sigma}_\varepsilon^2(\delta)$，我们有确切的估计值

$$\hat{\sigma}_\varepsilon^2(\delta) \equiv \frac{1}{N}\sum_{t=0}^{N-1}\frac{e_t^2}{v_t}. \tag{368}$$

代替式(363c)的递减对数似然函数，我们有

$$l(\delta \mid \boldsymbol{X}) \equiv N\log(\hat{\sigma}_\varepsilon^2(\delta)) + N\log\left(\frac{\Gamma(1-2\delta)}{\Gamma^2(1-\delta)}\right) + \sum_{t=1}^{N-1}(N-t)\log(1-\phi_{t,t}^2).$$

把式(368)带入上式，可以得到δ的函数，我们能数值极小化求出极大似然估计$\hat{\delta}$. 在有了$\hat{\delta}$之后，把$\hat{\delta}$代入式(368)可以得到σ_ε^2的极大似然估计$\hat{\sigma}_\varepsilon^2$.

9.4 平稳和非平稳分形差分过程的极大似然估计

本节，我们考虑分形差分过程形成参数δ和σ_ε^2的极大似然估计的另一种途径. 这里我们确定的思想是直接根据小波系数形成似然函数而不是通过时间序列本身. 这种方法有利于我们处理在确定性趋势情形观察的普遍分形差分过程；也就是不再需要对时间序列假设一个平稳模型(对于附加的细节参见 Craigmile 等 2000a，2000b).

相应地，对于给定的非负整数d，假设$\{U_t\}$是一个随机过程，它的d阶向后差分$\{Y_t\}$是具有参数$\delta^{(s)}$和σ_ε^2的零均值高斯平稳分形差分过程，其中$-\frac{1}{2}\leqslant\delta^{(s)}<\frac{1}{2}$，$\sigma_\varepsilon^2>0$. 有了这些基础，$\{U_t\}$是参数为$\delta=d+\delta^{(s)}$和$\sigma_\varepsilon^2$的一个分形差分过程，并且当$d=0$($d>0$)时是平稳的(不平稳的). 在假设我们实际观察到

$$X_t \equiv T_t + U_t,\quad t=0,1,\cdots,N-1,$$

实现的前提下，我们考虑估计δ和σ_ε^2的问题，其中T_t是未知的确定性趋势(Brillinger，1994 和 1996 考虑了一个和这个类似的模型，但是他提出一个关于趋势估计而不是此处所提到的不同的基于小波的方案). 为了简化，假设该趋势为r次的多项式：

$$T_t \equiv \sum_{j=0}^{r} a_j t^j$$

(正如在 Craigmile 等，2000b 中所讨论的，事实上，下面该方案是从纯多项式趋势来处理这种偏差的). 用向量记号，将上式写为$\boldsymbol{X}=\boldsymbol{T}+\boldsymbol{U}$.

令$\mathcal{W}$是基于一个宽为L的 Daubechies 小波滤波器的第J_0层的部分离散小波变换矩阵，$\boldsymbol{W}=\mathcal{W}\boldsymbol{X}$为相应的离散小波变换系数. 回忆这个滤波器涉及$L/2$阶的隐式差分操作，并且当$L/2\geqslant r+1$时，有把$r$阶多项式趋势减为零的重要性质. 这种性质实质表现在图 369 中，说明了当用哈尔、D(4)和 D(6)滤波器对线性

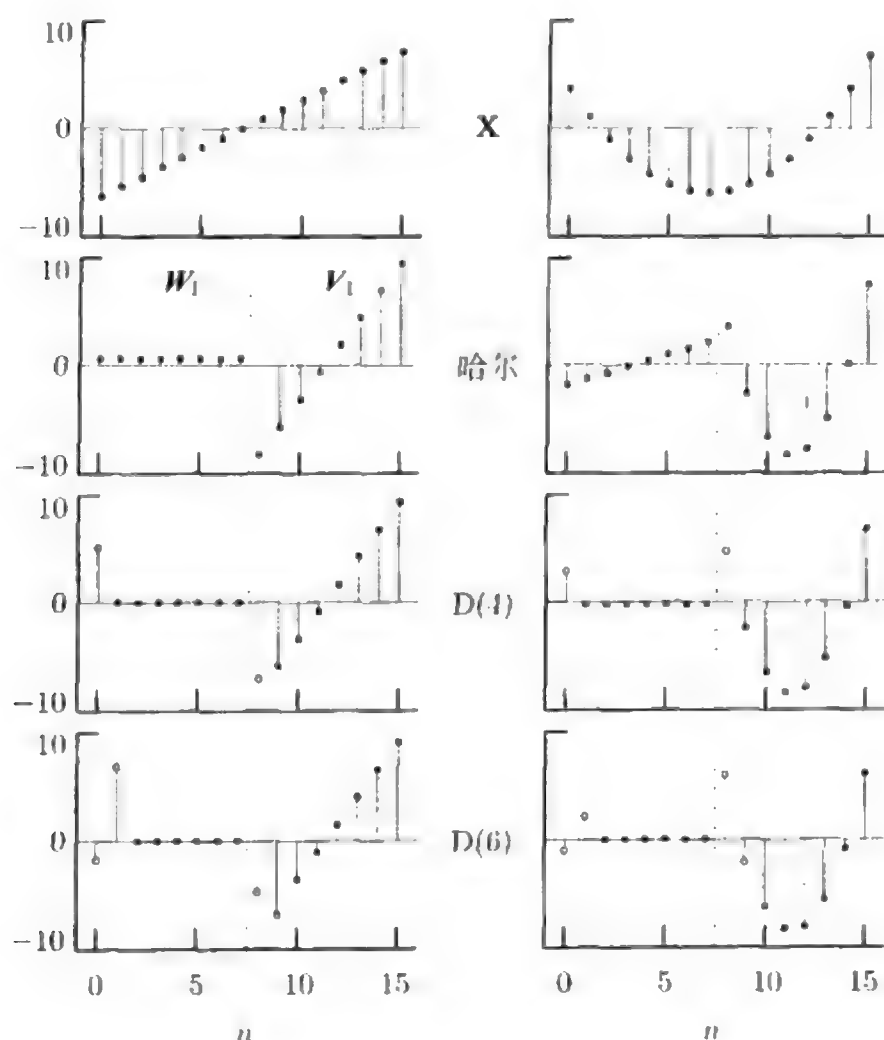


图 369 图中显示了线性和二次趋势 $\boldsymbol{T}$(上图)，基于哈尔、D(4)和 D(6)小波滤波器(分别是第 2 行至第 4 行)的部分离散小波变换系数 $\boldsymbol{W}=[\boldsymbol{W}_1^T, \boldsymbol{V}_1^T]^T$. 垂直虚线表示子向量 $\boldsymbol{W}_1$ 和 $\boldsymbol{V}_1$. 边界小波和尺度系数用圆圈表示(哈尔小波中没有，D(4)小波在每个子向量 $\boldsymbol{W}_1$ 和 $\boldsymbol{V}_1$ 中有一个，D(6)小波在每个子向量中有两个). 由于哈尔小波不能把线性或者二次多项式的趋势减为零，因此它的小波系数都是非零的；另一方面，由于 D(6)小波能把线性和二次多项式的趋势都减为零，因此它的 6 个非边界小波系数在两个 $\boldsymbol{W}_1$ 中都为零. 而 D(4)小波可以处理线性多项式却不能处理二次多项式，因此它的 7 个非边界小波系数在前者都为零，而后者都是非零的(作为资料，线性和二次趋势定义为 $T_t=0.9\cdot(t-7)$ 和 $T_t=0.2\cdot(t-7)^2-6$，这个注解来自 W. Constantine，MathSoft，Seattle)

$(r=1)$和二次$(r=2)$趋势计算时，一层离散小波变换系数 $\boldsymbol{W}_1$ 和 $\boldsymbol{V}_1$ 看起来类似. 用小波滤波器缩减为零的最高次多项式是 $r=\frac{L}{2}-1$，对于滤波器宽为 2，4 和 6，其分别是 0，1 和 2. 如果我们回忆在 $\boldsymbol{W}_1$(和 $\boldsymbol{V}_1$)中前 $L_1'=\frac{L}{2}-1$ 个系数被离散小波变换的周期边界条件所影响(参见式(146b))，则这个图说明了当$\frac{L}{2}\geqslant r+1$ 时，多项式趋势正好在边界小波系数和尺度系数中可以找到——非边界小波系数必定是

零并且没有被趋势所影响.

记住这些思想，让我们把离散小波变换系数 $\mathbf{W}$ 写成两个 N 维向量之和，即

$$\mathbf{W}=\mathbf{W}_{bs}+\mathbf{W}_{nb},$$

其中 $\mathbf{W}_{bs}$ 是将 $\mathbf{W}$ 除了边界小波系数和尺度系数之外的其他系数都置零得到的，而 $\mathbf{W}_{nb}$ 是将 $\mathbf{W}$ 除去非边界小波系数之外的其他系数都置零得到的. 由于 $\mathbf{X}=\mathcal{W}^{T}\mathbf{W}$，则有

$$\mathbf{X}=\mathcal{W}^{T}\mathbf{W}_{bs}+\mathcal{W}^{T}\mathbf{W}_{nb}\equiv\hat{\mathbf{T}}+\hat{\mathbf{U}},\tag{370a}$$

其中 $\hat{\mathbf{T}}$ 表示多项式趋势. 而 $\hat{\mathbf{U}}$ 包含变差，其源于估计趋势，该趋势服务于 $\mathbf{U}$ 的替代(这个分解例子在图 383 中进一步说明).

保存 $\mathbf{W}_{bs}$ 来处理 $\mathbf{T}$. 现在转到形成分形差分过程 $\{U_t\}$ 的参数 δ 和 σ_ε^2 的近似极大似然估计. 我们将 $\mathbf{W}_{nb}$ 中的非边界小波系数作为这些估计值的基础. 为此，必须确保小波滤波器选择有充分大的 L 以至于这些系数可作为零均值平稳过程的实现的一部分. 通过假设，$\{U_t\}$ 的 d 阶向后差分 $\{Y_t\}$ 是一个零均值平稳过程，因此练习[305]告诉我们必须挑选 $\frac{L}{2}\geqslant d$. 当结合条件 $L/2\geqslant r+1$ 时，为了恰当地处理 $\mathbf{T}$ 和 $\mathbf{U}$，要求 $L/2\geqslant\max\{r+1,d\}$(实际上，可以从 $\mathbf{X}$ 的图形和它的低阶差分获得合理的 L，例如，图 318 和 328). 正如上节所说的，这里我们集中精力在长记忆过程，但是我们所表述的理论对于 $\delta\in[-1/2,0]$ 的短记忆分形差分过程也是有意义的，由于 $\delta<d+\frac{1}{2}$，理论中宽为 L 的滤波器所处理的 δ 的范围是 $\left[-\frac{1}{2},\ \frac{L}{2}+\frac{1}{2}\right]$. 然而，事实上，在这个区间的末端估计 δ 有显著的偏差：当 $\delta\in[-1/2,-1/4]$ 时，哈尔和(较小程度上)D(4)滤波器是有问题的，并且在区间 $\delta\in\left[\frac{L}{2}+\frac{1}{4},\ \frac{L}{2}+\frac{1}{2}\right]$ 的上区域上，宽为 L 的滤波器有困难. 故对 δ 的建议区间为

$$\delta\in\left[-\frac{1}{4},\frac{L}{2}+\frac{1}{4}\right].\tag{370b}$$

如果从下面所描述的过程中获得的 δ 的估计大于 $\frac{L}{2}$，则比较而言，用宽为 $L+2$ 的小波滤波器计算这一估计是一种好的想法(类似地，如果我们用哈尔或者 D(4)小波得到一个低于零的估计，则应该把它和 D(6)估计进行比较).

在假设 $L/2\geqslant\max\{r+1,d\}$ 的前提下，第 j 层非边界小波系数由下式给出：

$$W_{j,t},\quad t=L_j',\cdots,N_j-1,\quad \text{其中}\quad L_j'\equiv\left\lceil(L-2)\left(1-\frac{1}{2^j}\right)\right\rceil,\tag{370c}$$

并且要求 $N_j'\equiv N_j-L_j'>0$(L_j' 的表示由式(146a)给出，并且在表 136 中对宽 $L=2,4,\cdots,20$ 的滤波器给予了计算). 上面的系数是具有谱密度函数的零均值高斯平稳过程的一部分，并且依赖于未知参数 δ 和 σ_ε^2，其中谱密度函数由

$\mathcal{H}_j(f)S_X(f)$给出. 令J_0是使$N'_{J_0}>0$的层(事实上,我们经常令J_0尽可能大;也就是固定它使$N'_{J_0}>0$而$N'_{J_0+1}\leqslant 0$). 定义$\boldsymbol{W}'_{nb}$是一个包含所有非边界小波系数的向量(回忆$\boldsymbol{W}_{nb}$的一些元素零化——去掉这些元素就得到$\boldsymbol{W}'_{nb}$). $\boldsymbol{W}'_{nb}$的长度是$N'\equiv\sum_{j=1}^{J_0}N'_j$,给定$\boldsymbol{W}'_{nb}$,$\delta$和$\sigma_\varepsilon^2$的精确似然函数为

$$L(\delta,\sigma_\varepsilon^2\mid\boldsymbol{W}'_{nb})\equiv\frac{1}{(2\pi)^{N'/2}\mid\Sigma_{\boldsymbol{W}'_{nb}}\mid^{1/2}}\mathrm{e}^{-[\boldsymbol{W}'_{nb}]^{\mathrm{T}}\Sigma^{-1}_{\boldsymbol{W}'_{nb}}\boldsymbol{W}'_{nb}/2}.$$

现在，我们要求通过处理$\boldsymbol{W}_{nb}$中的随机变量，尽管它们是无关的，利用离散小波变换的去相关性得到上式的一个有效近似. 由于高斯的假设，现在我们可以把似然函数近似地作为一元高斯密度之积，这仅不同于它们的联合的方差. 式(343a)对$\mathrm{var}\{W_{j,t}\}$给出一种确切表示，但是我们可以通过式(343c)中的C_j或者式(344)和式(354)中的$\widetilde{C}_j$来近似它(前者对$j\geqslant 3$有效，后者要求$j=1$或2，但是如果要求的话，对所有的j均可使用). 为了记号方便，我们采用C_j来近似，但是用$\mathrm{var}\{W_{j,t}\}$或者$\widetilde{C}_j$在下面代替C_j是件容易的事情，因此得到

$$L(\delta,\sigma_\varepsilon^2\mid\boldsymbol{W}'_{nb})\approx\widetilde{L}(\delta,\sigma_\varepsilon^2\mid\boldsymbol{W}'_{nb})\equiv\prod_{j=1}^{J_0}\prod_{t=0}^{N'_j-1}\frac{1}{(2\pi C_j)^{1/2}}\mathrm{e}^{-W^2_{j,t+L'_j}/(2C_j)}. \tag{371a}$$

练习[371] 用一种类似于产生成式(363b)和(363c)的方法，证明对于给定的δ，σ_ε^2的近似极大似然估计为

$$\tilde{\sigma}_\varepsilon^2(\delta)=\frac{1}{N'}\sum_{j=1}^{J_0}\frac{1}{C'_j(\delta)}\sum_{t=0}^{N'_j-1}W^2_{j,t+L'_j}. \tag{371b}$$

并证明缩减的对数似然函数是

$$\tilde{l}(\delta\mid\boldsymbol{W}'_{nb})\equiv\tilde{l}(\delta,\tilde{\sigma}_\varepsilon^2\mid\boldsymbol{W}'_{nb})-N'=N'\log(\tilde{\sigma}_\varepsilon^2(\delta))+\sum_{j=1}^{J_0}N'_j\log(C'_j(\delta)), \tag{371c}$$

把$\tilde{\sigma}_\varepsilon^2(\delta)$表达式代入上式可仅作为$\delta$的函数. ◁

对于δ和σ_ε^2求得的近似极大似然估计的方案如前是相同的：我们将数值地决定$\tilde{\delta}^{(s/ns)}$值. 该值最小化上面作为一个δ的函数，在$\delta=\tilde{\delta}^{(s/ns)}$时，计算$\tilde{\sigma}_\varepsilon^2(\delta)$得到对于$\sigma_\varepsilon^2$的极大似然估计$\tilde{\sigma}_\varepsilon^2\equiv\tilde{\sigma}_\varepsilon^2(\tilde{\delta}^{(s/ns)})$. Craigmile 等(2000a)证明了在假设$\boldsymbol{W}'_{nb}$中的小波系数确实不相关的情况下，对于大的N和$\delta\in[-1/2,L/2]$，估计值$\tilde{\delta}^{(s/ns)}$近似于均值为δ、方差为σ_δ^2的高斯分布，其中

$$\sigma_\delta^2\equiv 2\Big[\sum_{j=1}^{J_0}N_j'\gamma_j^2-\frac{1}{N'}\Big(\sum_{j=1}^{J_0}N_j'\gamma_j\Big)^2\Big]^{-1}, \tag{371d}$$

表 372

δ		MLE			
		哈尔	D(4)	LA(8)	确切值
0.4	均值	0.3670	0.3762	0.3792	0.3900
	偏差	−0.0330	−0.0238	−0.0208	−0.0100
	SD	0.0588	0.0732	0.0943	0.0437
	RMSE	0.0674	0.0769	0.0966	0.0448
	σ_δ	0.0530	0.0673	0.0869	
0.75	均值	0.7230	0.7277	0.7346	0.7677
	偏差	−0.0270	−0.0223	−0.0154	0.0177
	SD	0.0783	0.0878	0.0863	0.0272
	RMSE	0.0829	0.0906	0.0877	0.0325
	σ_δ	0.0526	0.0665	0.0857	

注：如同表 364，不同的是应用式(371a)的似然函数去定义对于 δ 的基于小波的近似极大似然估计 $\tilde{\delta}^{(s/ns)}$ ($\delta=0.4$ 时最后一列的精确极大似然估计与表 364 中的完全相同)．项 σ_δ 的性质在正文中说明．

而

$$\gamma_j \equiv \frac{\frac{\mathrm{d}\,\mathrm{var}\{W_{j,t}\}}{\mathrm{d}\delta}}{\mathrm{var}\{W_{j,t}\}} = -\frac{4\sigma_\varepsilon^2}{\mathrm{var}\{W_{j,t}\}}\int_0^{1/2} \mathcal{H}_j(f)\,\frac{\log(2\sin(\pi f))}{[2\sin(\pi f)]^{2\delta}}\mathrm{d}f. \qquad (372)$$

为了研究$\tilde{\delta}^{(s/ns)}$的统计性质，我们执行类似于上节对于样本大小 $N=256$ 所描述的一种模拟研究(参见表 364)，例外的是我们用 $\delta=0.75$ 代替 $\delta=0.25$(正如 7.8 节的评论与扩展所讨论的，通过累计求和模拟的 FD(−0.25)过程，我们可以模拟一个 FD(0.75)过程)．表 372 的前三列概括了哈尔、D(4)和LA(8)小波研究的结果．作为基准，最后一列给出了对于精确极大似然估计$\hat{\delta}$的对应统计($\delta=0.4$ 情况是从表 364 中拷贝而来，而 $\delta=0.75$ 情况是基于把它增加到 δ 的极大似然估计，在差分后得到的模拟 FD(0.75)序列)．当基于小波的近似极大似然估计具有平方误差均值的平方根，是精确极大似然估计的平方误差均值的平方根的两倍或三倍时，我们必须回忆下面的模型满足一个趋势存在，尽管在模拟研究中令 $\boldsymbol{T}=\boldsymbol{0}$——如果这种趋势存在，我们知道精确极大似然估计非常糟糕地执行．额外地，$\delta\geqslant\frac{1}{2}$的精确极大似然估计的使用要求原始时间序列的差分，反之当使用$\tilde{\delta}^{(s/ns)}$时，对于一个关于差分的预先决定是没有必要的．

表 372 也给出了 σ_δ，即式(371d)中对于$\tilde{\delta}^{(s/ns)}$的大的样本方差的平方根．我们用关于 $\mathrm{var}\{W_{j,t}\}$的式(343c)的带通近似 C_j 和式(372)积分的类似近似来计算 σ_δ．当 σ_δ 低估了所观察到的标准离差和平方误差均值的平方根(在 $\delta=0.75$ 时特别对于哈尔和 D(4)小波)时，我们在实际中用它作为$\tilde{\delta}^{(s/ns)}$中变化性的一种度量(另外的计算机模拟给出了 σ_δ 和观察到的平方误差均值的平方根随着 N 的增大而收敛)．

假设现在我们已经得到了时间序列的一个估计$\tilde{\delta}^{(s/ns)}$，并且想通过不存在趋势的零假设($\boldsymbol{T}=\boldsymbol{0}$)来接近假设时间序列中含有有意义的趋势成分．考虑 $P\equiv$

$\|\mathbf{X}\|^2/\|\mathbf{X}-\hat{\mathbf{T}}\|^2$为检验统计量．这个统计量是整时间序列的能量和非趋势序列或者非边界小波系数的能量之比，如下练习所示．

练习[373] 证明

$$P=\frac{\|\mathbf{X}\|^2}{\|\mathbf{W}_{nb}\|^2}.$$ ◁

注意由于 P 是能量比，因此它不依赖于 σ_ε^2 的值．在零假设和关于 $\tilde{\delta}^{(s/ns)}$ 条件下，我们可以通过重复模拟具有 $\sigma_\varepsilon^2=1$ 的零均值 FD($\tilde{\delta}^{(s/ns)}$)过程和计算每次模拟的 P 值来经验地确定 P 的分布．设 P_α 为模拟 P 值的上 100α%百分率点，特别地设 $\alpha=0.05$ 或者更小．当 $\mathbf{T}\neq\mathbf{0}$ 时，统计量 P 应该大，所以我们拒绝水平 α 上的零假设，如果 P 超出 P_α.

对这个实验需要注意两点．首先，虽然直观上的趋势是在大尺度上一个光滑非随机变分，如果时间序列具有显著的非零样本均值，实验的统计量 P 同样很大．为了防止拒绝零假设只不过由于是非零均值，我们调整检验统计量为

$$\bar{P}\equiv\frac{\sum_{t=0}^{N-1}(X_t-\bar{X})^2}{\|\mathbf{W}_{nb}\|^2},\tag{373}$$

随后将应用于时间序列和经验地确定零假设下的 $\bar{P}$ 分布的模拟序列上(由于小波系数不受样本均值的影响，所以 P 和 $\bar{P}$ 是一样的)．其次，我们可以通过对大量样本分布的模拟来补偿对观察估计 $\tilde{\delta}^{(s/ns)}$ 状态的影响．因此，预先产生每个模拟序列以得到 P 或者 $\bar{P}$ 的分布，从而产生偏离具有均值 $\tilde{\delta}^{(s/ns)}$ 的高斯分布和式(371d)确定的方差 $\delta=\tilde{\delta}^{(s/ns)}$ 的高斯分布．然后使用这个偏离作为模拟分形差分过程中的 δ 参数．

在原子钟数据上检验统计量 $\bar{P}$ 的一个应用的例子将在 9.7 节给出．

9.4 节的评论与扩展

[1]式(371a)的似然函数能够容易地适合于处理那些在大的尺度上可以用分形差分过程很好地建立模型的时间序列，但是对于小的尺度却不适合：对某个 $J_1>1$，如果模型需要拖延尺度 $j=J_1,\cdots,J_0$，只需将(371a)的结果的下限用 $j=J_1$ 来替换即可．这个趋势检验通过下面的过程同样容易调整：(i)用尺度系数 $\mathbf{V}_{J_1-1}$ 替换 $\mathbf{X}$；(ii)从 $\mathbf{W}_{nb}$ 中除去层 $j=1,\cdots,J_1-1$ 的尺度．

另外，如果趋势并不重要从而只对估计 δ 和 σ_ε^2 感兴趣的话，很容易调整式(371a)的似然函数使其也可以处理样本尺寸 N 不是 2 的整数幂的情况．为此，令 M 是使 $M>N$ 的 2 的最小次幂，由式(308c)通过在$\{X_t\}$的末端附加 $M-N$ 个零点来构造$\{X'_t\}$．令 $\mathbf{W}'_j$ 是$\{X'_t\}$的第 j 层小波系数．思想是，使用 $\mathbf{W}'_j$ 的元素形成似然函数，$\mathbf{W}'_j$ 的构造不包含 $\mathbf{X}'$ 端点的 $M-N$ 个零点中的任何一个，这些是

式(308d)中精确的 M_j' 系数，其中 $M_j' \equiv \left\lfloor \frac{N}{2^j} - 1 \right\rfloor - L_j' + 1$. 如果令 J_0 是满足 $M_{J_0}' > 0$ 的最大整数，那么式(371a)的似然函数就具有如下形式：

$$\prod_{j=1}^{J_0} \prod_{t=0}^{M_j'-1} \frac{1}{(2\pi C_j)^{1/2}} e^{-[W_{j,t+L_j'}']^2/(2C_j)},$$

其中，$W'_{j,t}$ 是 W_j' 的第 t 个元素.（这个过程的改善是利用在 Coifman and Donoho，1995，中讨论过的"循环旋转"这一观点，即循环平移$\{X_t'\}$，仅用那些不包含额外 $M-N$ 个零的小波系数来构造一个似然函数，从似然方程中决定 δ 的极大似然估计，然后平均极大似然估计获得所有可能的循环平移.）

[2]我们可以采用上面的方法来获得 δ 的一个"瞬时"估计量，它对追踪那些长记忆参数展开超时的非平稳长记忆过程很有用. 基本的想法是利用在同一时间的小波系数. 为了避免在时域上离散小波变换中二次抽样固有的歧义，可以在极大重叠离散小波变换的每个尺度舍掉一个小波系数用 $2^{j/2}$ 来乘它，然后就形成了一个由重新规范化的极大重叠离散小波变换系数组成的似然函数. 为了确保在时域上恰当地联合定位，很有必要利用最接近对称或 coiflet 滤波器——具有恰当的平移——因为这些滤波器近似地具有零相位，并且它们的输出可以在时域上关于原始时间序列做排列(详情参见 4.8 和 4.9 节). 这个输出结果简化了对数似然函数使得 δ 的"瞬时"估量在时间 t 处具有如下形式：

$$J_0 \log\left(\frac{1}{J_0}\sum_{j=1}^{J_0} \frac{2^j}{C_j'(\delta)} \widetilde{W}_{j,t_j}^2\right) + \sum_{j=1}^{J_0} \log(C_j'(\delta)), \tag{374}$$

其中 t_j 是第 j 层极大重叠离散小波变换系数的下标，它与时间 t 有关(上面的检验参见练习[9.6]).

9.5　分形差分过程的最小二乘估计

对一个长记忆过程$\{X_t\}$，在这一章的序言中我们记录了，绘尺度 τ_j 函数的小波方差 $\nu_X^2(\tau_j)$的 log/log 图，当 τ_j 变大时，这个图给出了一个近似的线性变化；即对某个 ζ 和 β 我们有 $\log(\nu_X^2(\tau_j)) \approx \zeta + \beta\log(\tau_j)$. 斜率 β 与近似幂规律谱密度函数的指数 α 有关，在低频处有$\beta = -\alpha - 1$. 现在，对 FD(δ)过程 $\alpha = -2\delta$(参见图 286 右上方的图)，所以近似地有 $\beta = 2\delta - 1$. 这个简单的联系暗示在尺度范围内我们能关于 β 的估计值形成 δ 的估计值，而用作为 $\log(\tau_j)$函数的对数小波方差得到(例如，参见 Abry 等，1993 和 1995，Abry and Veitch，1998，及 Jensen，1999a). 在这一节中，我们讨论如何建立一个 β 的加权最小二乘估计值——然后是 δ——考虑小波方差估计值的大样本性质(Abry 等，1995，和 Abry and Veitch，1998，采用了类似的方法).

让我们开始将小波方差与式(343b)的平均谱密度函数层 C_j 相联系，从中可以了解到当 τ_j 多大时线性逼近会相当好一些. 由定义可知小波方差 $\nu_X^2(\tau_j)$是极大重叠离散小波变换小波系数$\widetilde{W}_{j,t}$的方差. 离散小波变换和极大重叠离散小波变换

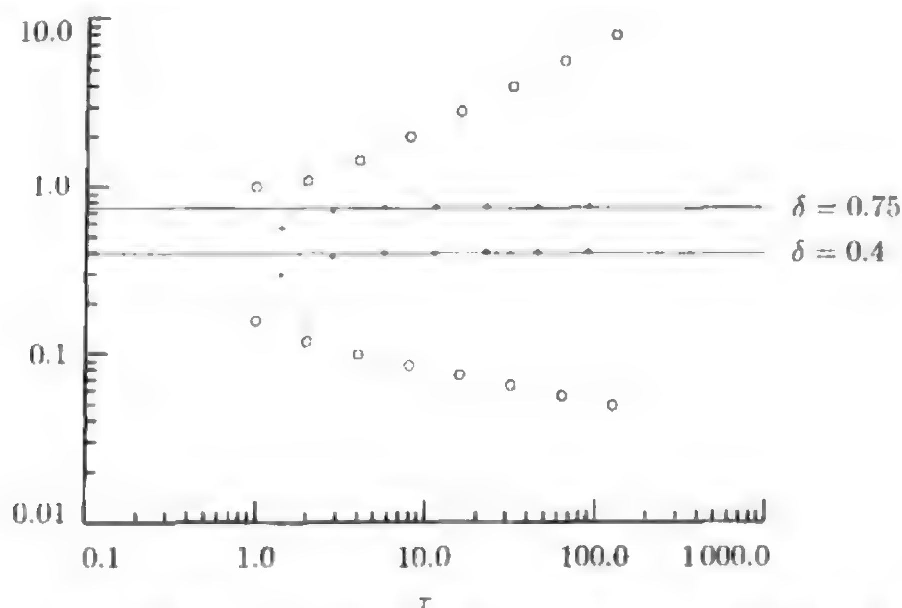


图 375 对于分形差分过程取 $\delta=0.4$(下面的圆圈)和 0.75(上面的圆圈)时，对作为 τ_j 的函数的小波方差 $\nu_X^2(\tau_j)$ 作数值积分(由式(343b)计算)的近似 $C_j/2^j$，$j=1,\cdots,8$. 这些圆圈集合中图的解释部分见正文

系数之间的联系，即 $W_{j,t}/2^{j/2}=\widetilde{W}_{j,2^j(t+1)-1}$，这个关系式告诉我们 $\nu_X^2(\tau_j)=\mathrm{var}\{W_{j,t}\}/2^j$. 式(343b)说明平均谱密度函数层 C_j 是 $\mathrm{var}\{W_{j,t}\}$的一个近似. 因此就有 $\nu_X^2(\tau_j)\approx C_j/2^j$，其中分形差分过程中的 C_j 可以通过计算式(343c)中的数值积分来确定.

图 375 在 log/log 坐标轴上给出了分形差分过程取 $\delta=0.4$(下面的圆圈)和 0.75(上面的圆圈)时作为 τ_j 的函数的 $C_j/2^j$，$j=1,\cdots,8$. 对于任何一个 δ，这些圆圈大约落在一条直线上，但是最左边的圆圈($j=1$)显著地向上偏离. 利用关系式 $\delta=\frac{1}{2}(\beta+1)$，当 τ_j 增大时，相邻两个圆圈之间的斜率可以确定所需的 δ. 令 δ_j 是由指标为 j 和 $j+1$ 的圆圈的斜率在 log/log 平面确定的：

$$\delta_j\equiv\frac{1}{2}\left(\frac{\log_{10}(C_{j+1}/2^{j+1})-\log_{10}(C_j/2^j)}{\log_{10}(\tau_{j+1})-\log_{10}(\tau_j)}+1\right).$$

图 375 给出了当 $\delta=0.4$(下面的加号)和 0.75(上面的加号)时，作为$(\tau_{j+1}+\tau_j)/2$ 的函数的 δ_j 的图形，$j=1,\cdots,7$. 我们看到当 j 增大时 δ_j 快速收敛于 δ(用一条水平线来标记). 这些图形暗示当 $j\geqslant 2$ 或 3 的时候线性逼近是可行的(参见 Jensen，1999a，正式证明了对于具有 $|\delta|<1/2$ 的分形差分过程，当 $j\to\infty$ 时 $\nu_X^2(\tau_j)\to C\tau_j^{2\delta-1}$，其中 $C=e^\zeta$).

从而，当 $j\geqslant J_1$ 时，我们把 $\zeta+\beta\log(\tau_j)$ 作为 $\log(\nu_X^2(\tau_j))$的近似值，其中 ζ 和 β 是待估计的未知参数，并且具有代表性地置 J_1 是 2 或 3. 假定我们用宽度为 L 的小波滤波器并且认为时间序列是具有 $\delta\in\left[-\frac{1}{4},\ \frac{L}{2}+\frac{1}{4}\right]$(如式(370b))的分形差分过程的一部分，$X_0,\cdots,X_{N-1}$，的实现. 然后在尺度$J_1\leqslant j\leqslant J_0$ 上能使用式(306b)的无偏极大重叠离散小波变换估计值$\hat{\nu}_X^2(\tau_j)$来估计 $\nu_X^2(\tau_j)$，其中 J_0 在给定样本尺寸 N 和滤波器宽度 L 限制的情况下尽可能地大. 由式(313a)得到

$$\log(\hat{\nu}_X^2(\tau_j))\stackrel{d}{=}\log(\chi^2_{\eta_j})+\log(\nu_X^2(\tau_j))-\log(\eta_j),$$

其中等价自由度(EDOF)η_j 依赖于尺度 τ_j，并且在实际应用中可以用 8.4 节中概述的三个方法之一得到. Bartlett and Kendall(1946)证明了 $E\{\log(\chi^2_{\eta_j})\}=\psi\left(\frac{\eta_j}{2}\right)+\log(2)$和 $\operatorname{var}\{\log(\chi^2_{\eta_j})\}=\psi'\left(\frac{\eta_j}{2}\right)$，其中 $\psi(\cdot)$和 $\psi'(\cdot)$是式(275)的 Γ_2 和 Γ_3 函数. 于是我们有

$$E\{\log(\hat{\nu}^2_X(\tau_j))\} = \psi\left(\frac{\eta_j}{2}\right)+\zeta+\beta\log(\tau_j)-\log\left(\frac{\eta_j}{2}\right).$$

如果我们令

$$Y(\tau_j)\equiv\log(\hat{\nu}^2_X(\tau_j))-\psi\left(\frac{\eta_j}{2}\right)+\log\left(\frac{\eta_j}{2}\right), \tag{376a}$$

则得到线性回归模型

$$Y(\tau_j)=\zeta+\beta\log(\tau_j)+e_j,$$

对此，误差项

$$e_j\equiv\log\left(\frac{\hat{\nu}^2_X(\tau_j)}{\nu^2_X(\tau_j)}\right)-\psi\left(\frac{\eta_j}{2}\right)+\log\left(\frac{\eta_j}{2}\right)$$

在分布上等于随机变量 $\log(\chi^2_{\eta_j})-\psi\left(\frac{\eta_j}{2}\right)-\log(2)$. 注意 e_j 有零均值和方差 $\psi'\left(\frac{\eta_j}{2}\right)$，并且只要 η_j 在 10 附近或大于 10 的时候近似于高斯分布(参见图 276). 我们进一步假设 e_j 是两两互不相关的，鉴于 9.1 节中的讨论，这个假设是合理的. 用向量的记法，可以写作

$$\boldsymbol{Y}=A\boldsymbol{b}+\boldsymbol{e}, \tag{376b}$$

其中 $\boldsymbol{Y}\equiv[Y(\tau_{J_1}),\cdots,Y(\tau_{J_0})]^{\mathrm{T}}$；$A$ 是一个$(J_0-J_1+1)\times 2$ 阶矩阵，它的第一列全是 1，第二列是 $\log(\tau_{J_1}),\cdots,\log(\tau J_0)$；$\boldsymbol{b}\equiv[\zeta,\beta]^{\mathrm{T}}$；并且 $\boldsymbol{e}\equiv[e_{J_1},\cdots,e_{J_0}]^{\mathrm{T}}$ 是一个具有零均值，并且协方差矩阵 Σ_e 是对角线元素为 $\psi'\left(\frac{\eta_{J_1}}{2}\right),\cdots,\psi'\left(\frac{\eta_{J_0}}{2}\right)$的对角矩阵的随机向量.

我们现在可以求助于线性最小二乘理论来得到 $\boldsymbol{b}$ 的加权最小二乘估计量，即

$$\hat{\boldsymbol{b}}^{(\mathrm{wls})}\equiv(A^{\mathrm{T}}\Sigma_e^{-1}A)^{-1}A^{\mathrm{T}}\Sigma_e^{-1}\boldsymbol{Y}$$

(例如，参见 Draper and Smith，1998). $\hat{\boldsymbol{b}}^{(\mathrm{wls})}$的第二个元素是想要 β 的加权最小二乘估计量，即

$$\hat{\beta}^{(\mathrm{wls})}=\frac{\sum w_j\sum w_j\log(\tau_j)Y(\tau_j)-\sum w_j\log(\tau_j)\sum w_jY(\tau_j)}{\sum w_j\sum w_j\log^2(\tau_j)-(\sum w_j\log(\tau_j))^2}, \tag{376c}$$

其中，$w_j\equiv1/\psi'\left(\frac{\eta_j}{2}\right)$，并且所有的和式是在 $j=J_1,\cdots,J_0$ 上进行的(上式的证明参见练习[9.5]的第一部分). 在模型后面的假设前提下，估计值$\hat{\boldsymbol{b}}^{(\mathrm{wls})}$具有平均值 $\boldsymbol{b}$ 和协方差

$$\Sigma_{\hat{\boldsymbol{b}}^{(\mathrm{wls})}}=(A^{\mathrm{T}}\Sigma_e^{-1}A)^{-1}.$$

表 377

δ		WLSE			MLE
		哈尔	D(4)	LA(8)	确切值
0.4	均值	0.3925	0.4006	0.4044	0.3900
	偏差	−0.0075	0.0006	0.0044	−0.0100
	SD	0.0715	0.0886	0.1185	0.0437
	RMSE	0.0719	0.0886	0.1186	0.0448
0.75	均值	0.7398	0.7443	0.7435	0.7677
	偏差	−0.0102	−0.0057	−0.0065	0.0177
	SD	0.0779	0.0877	0.1196	0.0272
	RMSE	0.0786	0.0879	0.1198	0.0325
	$\sqrt{\mathrm{var}\{\hat{\beta}^{(\mathrm{wls})}\}}$	0.0891	0.1145	0.1552	

注：如表 372，但是现在是用基于小波的加权最小二乘估计量$\hat{\delta}^{(\mathrm{wls})}$(最后一列是从表 372 中照搬过来的)．这些加权最小二乘估计量是基于小波方差的无偏极大重叠离散小波变换估计量$\hat{\nu}_X^2(\tau_j)$，小波方差假定有等价自由度$\eta_j=\max\{M_j/2^j, 1\}$．我们对这三种小波均设$J_1=2$．表中给出的结果是基于样本尺寸$N=256$的1024个模拟时间序列．

特别地，这个协方差矩阵的右下方元素告诉我们

$$\mathrm{var}\{\hat{\beta}^{(\mathrm{wls})}\}=\frac{\sum w_j}{\sum w_j\sum w_j\log^2(\tau_j)-(\sum w_j\log(\tau_j))^2} \tag{377}$$

(这是练习[9.5]的第二部分)．由于$\delta=\frac{1}{2}(\beta+1)$，我们可以设$\hat{\delta}^{(\mathrm{wls})}\equiv\frac{1}{2}(\hat{\beta}^{(\mathrm{wls})}+1)$为分形差分参数的相应估计值．记作$\mathrm{var}\{\hat{\delta}^{(\mathrm{wls})}\}=\frac{1}{4}\mathrm{var}\{\hat{\beta}^{(\mathrm{wls})}\}$．

为了了解到这个加权最小二乘估计量完成的怎么样，我们进行类似于上一节的模拟研究(参见表 372)，取$N=256$，$\delta=0.4$和0.75．为了简便，我们和式(314c)一致来确定未知的等价自由度；即设$\eta_j=\max\{M_j/2^j, 1\}$，像以前一样，其中$M_j\equiv N-L_j+1$是用来构造估量$\hat{\nu}_X^2(\tau_j)$的极大重叠离散小波变换系数的数目．我们看一下对$J_1$的三种选择，即取1，2，3，并且可以发现$J_1=2$的时候对样本尺寸及所有的三种小波都给出了最好的平方误差均值的平方根(鉴于图 375，这是合理的)．样本尺寸和滤波器宽度对哈尔、D(4)和LA(8)小波分别要求$J_0=8$，6和5．表 377 总结了模拟研究的结果．我们看到对于哈尔小波平方误差均值的平方根要比另外两个小波小一些，毫无疑问这是因为哈尔小波滤波器较小的宽度给了我们更多的可用尺度．注意大多数情形下的偏差都十分小，这就意味着平方误差均值的平方根和标准离差几乎是一样的．还可以发现对于给定的小波当$\delta=0.4$和0.75时平方误差均值的平方根十分类似(这个结果与式(377)中$\mathrm{var}\{\hat{\beta}^{(\mathrm{wls})}\}$的表达式有关，并且这里采用的这种决定等价自由度的方法不依赖于δ)．为了比较起见，表 377 的最下面一行给出了三种小波理论方差$\mathrm{var}\{\hat{\beta}^{(\mathrm{wls})}\}$的平方根．可以看到真实的平方误差均值的平方根比理论推导的稍微要小一点(可能是因为真实的和假设的等价自由度或误差序列e_j里的非零相关性之间不协调)．从实际应用的观点看，$\mathrm{var}\{\hat{\beta}^{(\mathrm{wls})}\}$似乎给$\hat{\beta}^{(\mathrm{wls})}$的可变性中的变量提供了一

表 378

δ		WLSE			MLE
		哈尔	D(4)	LA(8)	确切值
0.4	均值	0.388 1	0.399 4	0.400 0	0.390 0
	偏差	−0.011 9	−0.000 6	−0.000 0	−0.010 0
	SD	0.066 9	0.070 0	0.072 8	0.043 7
	RMSE	0.068 0	0.070 0	0.072 8	0.044 8
0.75	均值	0.727 8	0.743 2	0.744 2	0.767 7
	偏差	−0.022 2	−0.006 8	−0.005 8	0.017 7
	SD	0.073 9	0.071 8	0.072 6	0.027 2
	RMSE	0.077 2	0.072 1	0.072 8	0.032 5

注：类似于表 377，但是现在是在有反射边界的情形下，使用基于小波方差的无偏极大重叠离散小波变换估计量的加权最小二乘估计量$\hat{\delta}^{(\mathrm{wls})}$来做的.

个保守的上界. 最后，如果比较一下表 377 中给出的平方误差均值的平方根和表 372 给出的极大似然估计，可以发现当$\delta=0.4$时，除了 LA(8)小波之外，近似值极大似然估计都超过了加权最小二乘估计量，当$\delta=0.75$时情况正好相反.

9.5 节的评论与扩展

[1]Jensen(1999a)用常规最小二乘估计量(OLSE)来考虑δ的估计. 借助于式(376b)的回归模型来说(这有点不同于 Jensen 所使用的模型)，这个估计量由

$$\hat{\boldsymbol{b}}^{(\mathrm{ols})}\equiv(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{Y}$$

给出，并且具有均值$\boldsymbol{b}$和协方差$\Sigma_{\hat{\boldsymbol{b}}^{(\mathrm{ols})}}=(A^{\mathrm{T}}A)^{-1}(A^{\mathrm{T}}\Sigma_{\epsilon}A)(A^{\mathrm{T}}A)^{-1}$，从中我们可以求出$\mathrm{var}\{\hat{\beta}^{(\mathrm{ols})}\}$的表达式(详情参见 Draper and Smith，1998). 当常规最小二乘估计量和加权最小二乘估计量都是$\boldsymbol{b}$的无偏估计量时，理论上在下述意义下加权最小二乘估计量要比常规最小二乘估计量好，如果我们的模型假设都成立的话，应当有$\mathrm{var}\{\hat{\beta}^{(\mathrm{wls})}\}<\mathrm{var}\{\hat{\beta}^{(\mathrm{ols})}\}$. 如果我们用常规最小二乘估计量取代加权最小二乘估计量重复此节描述的模拟研究，就会发现对表 377 中的所有δ及小波的组合，对于常规最小二乘估计量的平方误差均值的平方根大约是对于加权最小二乘估计量的两倍大. 所以，我们推荐选择$\hat{\beta}^{(\mathrm{wls})}$作为最小二乘估计.

[2]注意 8.3 节的评论与扩展的第[3]条，Greenhall 等(1999)用下面的式子来定义小波方差的一个有偏极大重叠离散小波变换估计量

$$\tilde{\nu}_X^2(\tau_j)=\frac{1}{2N}\sum_{t=0}^{2N-1}\widetilde{W}_{j,t}^2$$

(参见式(306c))，其中$\widetilde{W}_{j,t}^2$变量现在取对于

$$X_0,X_1,\cdots,X_{N-2},X_{N-1},X_{N-1},X_{N-2},\cdots,X_1,X_0$$

的极大重叠离散小波变换的系数(参见关于式(140)的讨论). 它们的结果暗示，在大的尺度上，这个有偏估计量的统计性质要比无偏估计量$\hat{\nu}_X^2(\tau_j)$好. 为了了解

这个有偏估计量是否可以导致 δ 的一个改进的加权最小二乘估计量，让我们构造一个估计量，记为$\tilde{\delta}^{(\mathrm{wls})}$，像式(376c)的右边那样定义它，但是现在要用$\tilde{\nu}_X^2(\tau_j)$来代替$\hat{\nu}_X^2(\tau_j)$构成 $Y(\tau_j)$. 有偏估计量的一个直接优势就是，对任一小波我们都可以令 $J_0=\log_2(N)$. 为了$\tilde{\delta}^{(\mathrm{wls})}$分布的进一步研究，我们假设基于 $\eta_j=N/2^j$ 的权 w_j. 利用这个设置，我们重新来做模拟研究并获得表 378 中的结果，这揭示了对于所有 δ 的组合及表 377 中提到的小波，平方误差均值的平方根都有改进. D(4) 和 LA(8)小波表现出很大的改进并且似乎比 $\delta=0.75$ 时的哈尔小波的改进还要好一点. 然而，仍然需要进一步的研究来确定$\hat{\delta}^{(\mathrm{wls})}$和$\tilde{\delta}^{(\mathrm{wls})}$各自的优缺点(特别地，当前者对于趋势的某一类型很好时，后者却不一定).

[3]像前一节讨论极大似然估计那样(参见评论与扩展的第[2]条)，我们可以采用这一节描述的估计量来获得 δ 的一个"瞬时"估计量. 通过对每个 j 层用一个单个的平方极大重叠离散小波变换系数来形成小波方差估计量，被选择的系数要满足在时间上是共同定位的. δ 在 t 时刻得到的估计量由

$$\frac{1}{2}\left(\frac{(J_0-J_1+1)\sum\log(\tau_j)Y_t(\tau_j)-\sum\log(\tau_j)\sum Y_t(\tau_j)}{(J_0-J_1+1)\sum\log^2(\tau_j)-(\sum\log(\tau_j))^2}+1\right) \qquad (379)$$

给出，其中 $Y_t(\tau_j)\equiv\log(\widetilde{W}^2_{j,t_j})-\psi\left(\frac{1}{2}\right)-\log(2)$；$t_j$ 与式(374)中的相同；像以前一样，所有的和式都是从 $j=J_1$ 加到 J_0(上面的验证作为练习[9.6]). 最小二乘和极大似然的优缺点至今还是未知.

[4]Geweke and Porter-Hudak(1983)，Robinson(1995)和 Hurvich 等(1998)讨论了长记忆参数的最小二乘估计量的统计性质，这些长记忆参数是基于回归作为对数频率的函数的对数周期. McCoy 等(1998)认为加权最小二乘估计量是基于对数多锥谱密度函数估计量(这比基于周期图的估计量要好). 当基于傅里叶和基于小波的最小二乘估计量之间确定的比较结果还没有给出的时候，Jensen (1999a)发现基于周期图的估计量不如他明确给出的常规最小二乘估计量好，而 McCoy and Walden(1996)却指出基于小波的$\hat{\delta}^{(s)}$比基于周期图和多锥的 δ 的估计量都要好.

9.6 方差的齐次检验

本节，我们考虑具有长记忆结构的时间序列的方差齐次检验的问题. 在 7.6 节，我们回顾了具有长记忆性质的随机模型，这些模型都要求平稳性假设. 这个假设说明这一过程的方差与时间无关. 因为长记忆过程的实现能够出现增长变化性的脉冲或相对静止的周期，所以当我们遇到一个使这个假设出现疑问的时间序列时，做一个统计检验来帮助评估齐次假设的合理性是很重要的. 一个明显的检验是将时间序列分割为子序列，计算每个子序列的样本方差，然后确定这些样本方差的差异是否与我们期望从平稳的长记忆过程中能观察到的相关；然而，就像

我们在8.1节中注意到的那样，样本方差的采样性质对于长记忆过程是有疑问的，所以我们仍然需要仔细研究子序列的选择方案如何对检验结果产生影响. 正如我们在本节表明的那样，可以借助于时域局部化系数，使用离散小波变换去长记忆过程相关这一事实来引出方差的齐次检验(详情参见 Whitcher，1998，和 Whitcher 等，2000a).

令 $\mathbf{X}\equiv[X_0, \cdots, X_{N-1}]^T$ 是我们计划用平稳的高斯分形差分过程建立模型的时间序列，但是方差的齐次假设尚有疑问. 为了方便，我们假定对某个 J_0，N 是 2^{J_0} 的整数倍(这个限制容易放宽——参见下面的评论与扩展中的第[1]条). 关键的思想是假设 $\mathbf{X}$ 的方差非齐次使得在某些选定的尺度上，它也能表示它本身. 从而，令 $\mathbf{W}_j$ 是 $\mathbf{X}$ 的离散小波变换的第 j 层小波系数. 我们假设 j 层满足，像式(370c)显示出的那样，存在 $N_j'>0$ 个非边界小波系数 $W_{j,L_j'}, \cdots, W_{j,N_j-1}$，其中像以前一样，$N_j'\equiv N_j-L_j'$. 我们将用这些小波系数对于零假设

$$H_0: \operatorname{var}\{W_{j,L_j'}\} = \operatorname{var}\{W_{j,L_j'+1}\} = \cdots = \operatorname{var}\{W_{j,N_j-1}\}$$

构造一个检验统计量. 我们假定在这里离散小波变换的去相关性有效，所以在假设 H_0 下，这些非边界系数可以认为是由零均值的高斯白噪声的一个样本(在 Whitcher，1998，和 Whitcher 等，2000a，中的模拟研究表明了去相关在这里给出的检验统计量方面是很好的). Brown 等(1975)，Hsu(1977)，Inclán and Tiao (1994)以及另外一些人研究一个可以区别 H_0 和各式各样的替换假设，包含

$$\operatorname{var}\{W_{j,L_j'}\} = \cdots = \operatorname{var}\{W_{j,t'}\} \neq \operatorname{var}\{W_{j,t'+1}\} = \cdots = \operatorname{var}\{W_{j,N_j-1}\} \tag{380a}$$

的检验统计量 D，其中 t'是一个未知的变化点. 这个检验统计量是基于正规化的累积平方和，并且在文中可以如下定义. 令

$$\mathcal{P}_k \equiv \frac{\sum_{t=L_j'}^{k} W_{j,t}^2}{\sum_{t=L_j'}^{N_j-1} W_{j,t}^2} \quad k = L_j', \cdots, N_j-2, \tag{380b}$$

并且定义 $D\equiv\max\{D^+, D^-\}$，其中

$$D^+ \equiv \max_k\left(\frac{k-L_j'+1}{N_j'-1} - \mathcal{P}_k\right), \quad D^- \equiv \max_k\left(\mathcal{P}_k - \frac{k-L_j'}{N_j'-1}\right).$$

注意在假设 H_0 下，$\mathcal{P}_k$ 的分子期望值与其分母期望值的比恰好是$(k-L_j'+1)/N_j'$，它随着 k 线性变化. 实质上，检验统计量度量了在假设 H_0 下所期望线性增加的最大偏差(正的或负的)的大小. 当 H_0 不成立时，则所期望的线性增加有偏差，我们可以希望 D 大到适合某个替换的假设，包含在式(380a)中出现的. 当 D"太大"时，我们将舍弃 H_0.

为了量化当 D“太大”时的情形，然后评估对给定的一个时间序列检验统计量是提供还是否定零假设，我们需要知道 D 在 H_0 下的分布．令 $d(p)$ 是作为这一分布的 $p\times 100\%$ 百分点；即 $\mathbf{P}[D\leqslant d(p)]=p$．对任意的 N_j'，还没有已知的容易处理的解析式来决定 $d(p)$，但幸运的是有两个很好的方法可以逼近它．首先，当 N_j' 大约是 128 或更大时，我们可以用 Inclán and Tiao(1994)给出的一个大的样本结果，即

$$\mathbf{P}[D>d(p)]=1-p\approx 2\sum_{l=1}^{\infty}(-1)^{l+1}\mathrm{e}^{-l^2 N_j' d^2(p)} \tag{381}$$

(右边是基于 Billingsley，1968，中的式(11.39))．为了看到在实际中如何运用这一表达式，假定我们想要我们的检验有显著性水平 α；即如果我们在一个抽样设置下重复地运用这个检验，α 是 H_0 实际上是正确的时候我们却(错误地)舍弃它的次数比(典型地，α 取小的值，如像 0.05 或 0.01)．然后，用检验统计的观察值来替换(381)右边式子中的 $d(p)$，并且评估得到的表达式　　如果表达式小于(大于)α，则我们舍弃(不舍弃)在显著性水平 α 的零假设．

逼近 D 的分布的第二种方法是进行模拟研究，从中获得偏离高斯随机数发生器的 N_j'，计算 D，然后多次重复这个过程直到建立了 $\mathbf{P}[D\leqslant d(p)]$的经验逼近．特别地，为了估计对于水平 α 的检验观察到的 D，假定生成 M 个高斯白噪声序列，并且这些序列中的 M' 个得到的 D 超过真实的小波系数的 D．如果 $M'/M<\alpha$ 我们将舍弃零假设．这里需要注意，M 要足够大以使得观察的比例 M'/M 的变化与它的 α 距离相比很小．近似的经验法则验证

$$\left(\frac{M'}{M}-\alpha\right)^2>\frac{4M'(M-M')}{M^3};$$

如果这个不等式不满足，就需要加大 M 以使不等式成立(当 $\alpha=0.05$ 或 0.01 时，如果加大 M 到超过 100 000 仍然不能使不等式成立，则检验的显著性显然很接近于 α，所以我们不能正式地舍弃或接受 H_0)．

于是，我们可以用检验统计量 D 来给出方差齐次假设的尺度对尺度的评估(9.8 节给出了一个例子)．假定在某一 j 层，我们可以利用 D 舍弃 H_0，并且能合理地接受式(380a)中出现的替换的假设；即在未知变化点有单一的可变性的变化．通过考虑在这个替换的假设下 $\mathcal{P}_k$ 分子的期望值与其分母期望值的比，我们可以讨论事实上生成 D 的指标 k 应当与变动点 t' 很接近．这个事实暗示应该令最大的指标作为 t' 的估计量；然而，由于识别一个时间序列本身的变动点更有意义，并且离散小波变换中固有的子抽样必然地在时域引入歧义，我们考虑用极大重叠离散小波变换来替换．在这一方案下，我们形成了 D 统计量的改进的方案，用非边界极大重叠离散小波变换小波系数来替换式(380b)中的离散小波变换小波

系数. 于是，有

$$\widetilde{\mathcal{P}}_k \equiv \frac{\sum_{t=L_j-1}^{k} \widetilde{W}_{j,t}^2}{\sum_{t=L_j-1}^{N-1} \widetilde{W}_{j,t}^2}, \quad k = L_j - 1, \cdots, N-2$$

和

$$\Delta_k^+ \equiv \frac{k - L_j + 2}{N - L_j} - \widetilde{\mathcal{P}}_k, \quad \Delta_k^- \equiv \widetilde{\mathcal{P}}_k - \frac{k - L_j + 1}{N - L_j}, \tag{382}$$

定义$\widetilde{D}^+ \equiv \max_k\{\Delta_k^+\}$，$\widetilde{D}^- \equiv \max_k\{\Delta_k^-\}$，$\widetilde{D} \equiv \max\{\widetilde{D}^+, \widetilde{D}^-\}$. 我们对变动点的评估是指标$\hat{k}$，得到辅助统计量$\widetilde{D}$；即有$\widetilde{D} = \Delta_{\hat{k}}^+$或$\widetilde{D} = \Delta_{\hat{k}}^-$. 如果我们用 Daubechies 最接近对称或 coiflet 小波滤波器来构造$\widetilde{D}$，就可以将$\hat{k}$映射到物理意义上的时间$t_0 + (\hat{k} + |\nu_j^{(H)}| \bmod N)\Delta t$，其中$t_0$是与$X_0$联系的实际时间；$\nu_j^{(H)}$是由对于最接近对称和 coiflet 滤波器的式(114c)和(124)各自给出的；而Δt是观测之间的抽样时间. 在 Whitcher(1998)和 Whitcher 等(2000a)中的模拟研究说明$\hat{k}$在时间序列本身中是变动点的有效估计量，随着变动过程前后方差的比值远离单位 1，实际变动点$\hat{k}$的分布变得更紧. 这些研究还指出，如果我们利用基于离散小波变换的统计量 D 在某些尺度上舍弃零假设，则像这样的最小尺度应当用于通过相应的$\widetilde{D}$来估计变动点(对于尼罗河序列 9.8 节给出了一个变动点估计的例子).

9.6 节的评论与扩展

[1]虽然在本节基于离散小波变换的检验统计量 D 的公式假定样本尺寸 N 是 2 的幂的倍数，实际上我们可以用同样的办法来处理任意样本的尺寸，就像 9.4 节的评论与扩展第[1]条中最后描述的那样；即我们可以在时间序列的末端加零，使得加零后的序列具有用标准离散小波变换容易处理的长度，然后仅用不包含这些零的 M_j' 个小波系数就可以形成 D(当形成基于极大重叠离散小波变换的辅助统计量$\widetilde{D}$时没有必要做类似的调整，因为极大重叠离散小波变换能自然地处理任意样本尺寸).

[2]注意辅助统计量$\widetilde{D}$与式(189)定义的旋转累积方差图之间的联系(对子潮海平面变化，图 190 给出了这样图的一个例子). 因为$\|\widetilde{\mathbf{W}}_j\|^2$是由$\|\mathbf{X}\|^2$在尺度$\tau_j$处的变化引起的，且如果循环移动$\widetilde{\mathbf{W}}_j$以获得合适的时间调整则$\|\widetilde{\mathbf{W}}_j\|^2$保持不变，这个图本质上与积累的循环移动$\widetilde{W}_{j,t}^2$和一致增大至$\|\widetilde{\mathbf{W}}_j\|^2$之间的差分成比例. 统计量$\widetilde{D}$与这些绝对差分的最大值的规范方案密切相关. 特别地，我们对应于物理意义上的时间 k 绘图 Δ_k^-，则得到的图非常类似于旋转累计方差图(参见下面对于一个例子的图 388)；然而，注意到 Δ_k^- 仅包含非边界极大重叠离散小波变换系数，但是旋转累计方差图用了所有的极大重叠离散小波变换系数(后者可以证明适合于像海平面变化这样的序列，经过仔细选择样本尺寸，就可以把它当作真正的循环来对待).

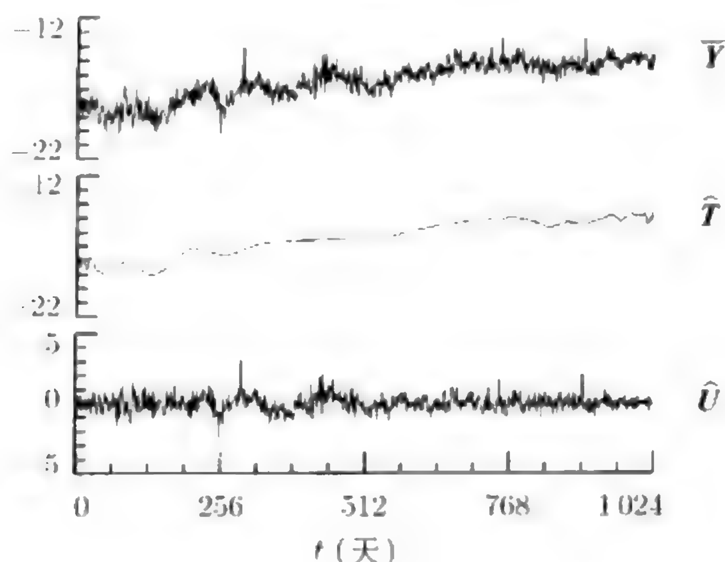


图 383　原子钟分形频率基于小波的分解偏离 $\bar{Y}$ 分为估计趋势$\hat{\boldsymbol{T}}$和这个趋势的残余$\hat{\boldsymbol{U}}$. 这里我们使用 $J_0=7$ 层的 LA(8)部分离散小波变换. 注意$\hat{\boldsymbol{T}}$比一个低阶多项式有更多的结构，并且实际上它类似于从一个可变带宽平滑器得到的输出：它在接近于序列的中间值很光滑，但是在趋向末端点的时候变得很粗糙

9.7　例子：原子钟偏差

现在，让我们用原子钟分形频率偏离$\{\bar{Y}_t(\tau_1)\}$来举例说明这一章的一些思想，我们先前在 1.2、1.3 和 8.6 节看到过. 这些偏离是在式(321c)中借助于测量$\{X_t\}$时间差分被定义，并且在图 318 中间绘出(这幅图有两个垂直的轴——右边的轴是对$\{\bar{Y}_t(\tau_1)\}$适用的轴).

就像在 8.6 节末尾注释的那样，对图 322 中$\{\bar{Y}_t(\tau_1)\}$的小波方差估计值的平方根与长记忆的解释一样，但是问题在于这些偏离是否就是通过平稳的或不平稳的过程建立的最好的模型，这个问题由于线性趋势的可能出现而变得混淆不清. 正因为这个，让我们首先接受模型$\bar{\boldsymbol{Y}}=\boldsymbol{T}+\boldsymbol{U}$(像 9.4 节那样)，其中 $\bar{\boldsymbol{Y}}$ 是一个包括$\bar{Y}_0(\tau_1)$，…，$\bar{Y}_{1023}(\tau_1)$的向量，并且 $\boldsymbol{T}$ 和 $\boldsymbol{U}$ 分量分别是多项式趋势和具有未知参数 δ 和σ_ε^2的零均值分形差分过程. 作为估计的式(370a)附加分解 $\bar{\boldsymbol{Y}}=\hat{T}+\hat{\boldsymbol{U}}$的一个例子，图 383 展示了 $J_0=7$ 层的 LA(8)离散小波变换处理(练习[9.7]将利用 D(4)和 C(6)小波产生一个类似的分解).

求式(371c)的极小化缩简的对数似然函数，我们得到 δ 的极大似然估计$\tilde{\delta}^{(s/ns)}$，然后利用式(371b)来获得 σ_ε^2的相应估计量，即$\tilde{\sigma}_\varepsilon^2\equiv\tilde{\sigma}_\varepsilon^2(\tilde{\delta}^{(s/ns)})$. 对于哈尔、D(4)和 LA(8)小波，表 384 的上面部分列出了这些估计量(还给出了基于式(371d)对$\tilde{\delta}^{(s/ns)}$的大样本标准偏离 $\sigma_{\tilde{\delta}^{(s/ns)}}$，并且在表的底部给出了适合每个小波的 J_0 的值). 虽然 δ 的 D(4)和 LA(8)极大似然估计很好地吻合并且说明平稳的分形差分模型是合适的，哈尔估计是相当大的并且几乎超出了平稳区域；然而，这个估计可能还不是很合适，因为哈尔小波甚至不能处理线性趋势.

表 384

	哈尔	D(4)	LA(8)
$\tilde{\delta}^{(s/ns)}$	0.503 1	0.394 3	0.392 1
$\sigma_{\tilde{\delta}^{(s/ns)}}$	0.025 2	0.028 2	0.031 8
$\tilde{\sigma}_\varepsilon^2$	0.305 7	0.283 8	0.274 0
$\hat{\delta}^{(wls)}$	0.444 9	0.381 2	0.346 0
$\sqrt{\mathrm{var}\{\hat{\delta}^{(wls)}\}}$	0.037 4	0.041 8	0.047 9
J_0	10	8	7

注：原子分形频率偏离的参数估计(详情参见正文).

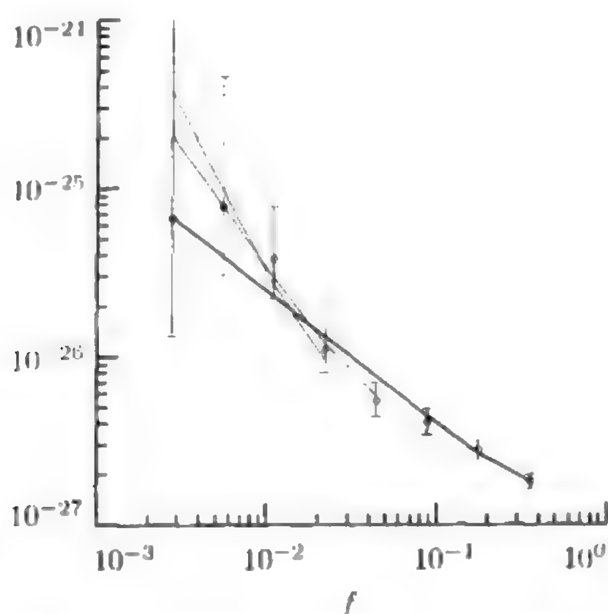


图 384 原子钟分形频率偏离的 D(4)极大重叠离散小波变换小波方差估计(再次表示为谱水平 C_j)，同时从 δ 和 σ_ε^2 的极大似然估计推算出谱水平(粗曲线). 那两条短的细的曲线将在正文中讨论

在 D(4)极大似然估计$\tilde{\delta}^{(s/ns)}$的条件下，现在可以检验无趋势假设，即在模型 $\bar{Y}=T+U$ 中 $T=0$. 从图 318 中很明显可以看出，分数频率偏离有一个显著的非零均值. 因为我们不希望在这个趋势中考虑这个非零均值，所以就要利用式(373)的检验统计量 $\bar{P}$. 对于观察值 $\bar{Y}$，我们得到 $\bar{P}\doteq 4.581$. 为了在零假设下估计 $\bar{P}$，我们用 Davies-Harte 算法(参见 7.8 节)来模拟长度 $N=1\ 024$ 的 10 000 个 FD(0.394 3)时间序列. 对于 10 000 个模拟序列 $\bar{P}$ 统计量的最小值，中间值及最大值分别为 1.019，1.182 及 2.047，所以我们可以有根据地拒绝零假设.

然而，让我们回头考虑一下从 D(4)小波得到的分形差分模型是否真的适合 $\bar{Y}$. 为此，我们取 D(4)估计量$\hat{\nu}_{\bar{Y}}(\tau_j)$(在图 322 中用圆圈表示)，利用关系式$C_j\approx 2^j\nu_{\bar{Y}}^2(\tau_j)$将它们(连同它们的联合 95%置信区间)转换为谱水平，然后绘出它们对

应于 $1/2^{j+\frac{1}{2}}$ 周期/天的图（这是在一个对数尺度上的倍频程带 $\left[\frac{1}{2^{j+1}\Delta t}, \frac{1}{2^j\Delta t}\right]$ 的中心频率）. 结果在图 384 中用圆圈表示，穿过它们的垂直线表示置信区间. 图中还有一条粗的曲线，这表示对应于 D(4) 极大似然估计 $\tilde{\delta}^{(s/ns)}$ 和 $\tilde{\sigma}_\varepsilon^2$ 的谱水平 C_j（这些水平是从式(343c)中利用数值积分计算出来的）. 当除了一个 95% 置信区间之外的其他区间限制了 D(4) 谱水平的时候，总体上来说吻合是有问题的，并且要求 $\bar{\mathbf{Y}}$ 应当比用简单的两参数分形差分模型能合理解决的情况有更多的结构.

根据 Jensen(1999b，2000)的思想，我们采用估计的方法来处理自回归分形整合移动平均模型，然后对 $\bar{\mathbf{Y}}$ 接受一种更复杂的模型；然而，鉴于我们所关心的是分形差分模型的适当性为了再评价趋势，我们考虑一种比较简单的方法. 就像图 384 中的圆圈所指出的那样，3 个或 4 个最小的尺度（即高频）看起来好像服从线性关系（在 log/log 空间），但是在一些较大的尺度上可以论证其与线性关系是不协调的. 我们可以很容易地采用上面的方法，对于 4 或 5 个最大的尺度，δ 的极大似然估计仅仅需要用非边界小波系数，并且接下来的趋势估计只需用这个尺度限制的极大似然估计就可以了. 如果令 J_1 是我们决定保留的最小尺度的指标，我们改进的方法是取 $\bar{\mathbf{Y}}$ 的 J_1-1 层部分离散小波变换且用尺度系数 $\mathbf{V}_{J_1-1}$ 作为我们拟合分形差分模型的序列. 这一方法的基本原理是只有 $\mathbf{V}_{J_1-1}$ 和 $\bar{\mathbf{Y}}$ 之间的差分是在 $j=1,\cdots,J_1-1$ 层的小波系数中，而当拟合分形差分模型的时候我们却想要忽略这些系数. 除了有限个小尺度边界小波系数之外，在 $\mathbf{V}_{J_1-1}$ 中也有 $\bar{\mathbf{Y}}$ 中的趋势，所以我们可以通过像分形差分过程那样建立模型 $\mathbf{V}_{J_1-1}$ 来估计趋势.

现在让我们将这个改进的方法应用到分形频率偏移，先取 $J_1=4$，再取 $J_1=5$. 当 $J_1=4$ 时，利用 $\mathbf{V}_3$ 得到 $\tilde{\delta}^{(s/ns)}\doteq 0.7025$ 的 D(4) 极大似然估计. 在图 384 中两条细曲线中较长的那条表示这个拟合模型给出的谱水平. 相应的观察检验统计量是 $\tilde{P}\doteq 8.113$. 当后者与 10 000 个模拟 FD(0.702 5) 序列（每个长度 $N_3=128$）的相似统计量相比较时，我们发现 1.74% 的模拟的 $\tilde{P}$ 都超过了观察到的 $\tilde{P}$，于是我们拒绝显著性水平 $\alpha=0.05$ 的零假设，而不拒绝显著性水平 $\alpha=0.01$ 的零假设. 另一方面，当 $J_1=5$ 时我们用 $\mathbf{V}_4$ 且仅保留 $\bar{\mathbf{Y}}$ 中最大的 4 个尺度（与图 322 给出的区分保持一致），我们得到 $\tilde{\delta}^{(s/ns)}\doteq 0.9978$（即很接近一个随机游动模型）. 相应的谱水平在图 384 中用较短的细曲线表示. 现在我们得到 $\tilde{P}\doteq 9.697$，这被 10 000 个模拟序列的相应统计量足足超过了 32% 的时间——这里我们不能拒绝任何合理显著性水平上的无趋势零假设！这个例子指出，时间序列趋势显著性的评估可以高度依赖于假定的统计模型，特别是在处理非平稳长记忆过程的时候.

为了得到表 384 中给出的与极大似然估计的比较结果，让我们也来计算相应的加权最小二乘估计量 $\hat{\delta}^{(\mathrm{wls})}$. 就像在 9.5 节讨论的那样，现在我们在尺度 $j=J_1,\cdots,J_0$ 上用非边界极大重叠离散小波变换系数，取 $J_1=2$. 得到的加权最小二乘估计在表 384 的中间部分给出，还有基于式(377)的标准误差估计及关系式 $\operatorname{var}\{\hat{\delta}^{(\mathrm{wls})}\}=\frac{1}{4}\operatorname{var}\{\hat{\beta}^{(\mathrm{wls})}\}$. 我们发现加权最小二乘估计量可与相应的极大似然估

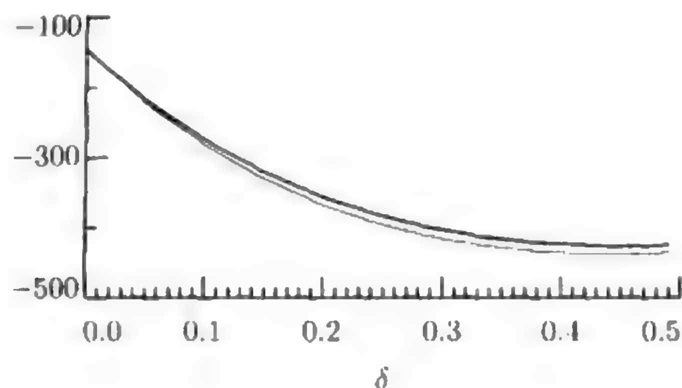


图 386 使用尼罗河时间序列(参见图 192)的最后 512 个观察值计算得到的分形差分参数 δ 的缩简对数似然函数. 粗的曲线是基于 LA(8)小波的并且在估计值$\tilde{\delta}^{(s)} \doteq 0.4532$ 时达到最小值，而细的曲线是针对真实方法的并且在$\hat{\delta} \doteq 0.4452$ 时达到最小值

计做比较——稍微小一点，而标准误差却稍微大一点.

9.8 例子：尼罗河最低水位

让现在我们用尼罗河最低水位(这个序列在图 388 的顶端重新绘出)来以实例说明 9.3 和 9.6 节中的方法. 像 5.9 和 8.8 节中讨论的那样，这个时间序列的大约前 100 年似乎与余下的序列有不同的小尺度变化(可能是由于 715 年左右测量工具的改变). 从而，我们沿用 Beran(1994)的方法，对除去第一部分的尼罗河时间序列建模形成具有未知过程均值的平稳分形差分过程. 因为总共有 663 个值且 9.3 节中的方法假定了一个具有 2 的幂次大小的样本，对这个例子来说仅用最后 $N=512$ 年(即 773 至 1284 年)是很方便的. 我们用最后 512 数据的样本均值 $\overline{X} \doteq 11.553$ 来估计过程均值. 使用 LA(8)小波，对 δ 和 σ_ε^2 的极大似然估计是$\tilde{\delta}^{(s)} \doteq 0.4532$ 及$\tilde{\sigma}_\varepsilon^2(\tilde{\delta}^{(s)}) \doteq 0.4293$，而相应的精确的极大似然估计是$\hat{\delta} \doteq 0.4452$ 及$\hat{\sigma}_\varepsilon^2(\hat{\delta}) \doteq 0.4230$(哈尔和 D(4)估计量的计算是练习[9.8]的主题). 决定 δ 每个极大似然估计的缩简的对数似然函数在图 386 中绘出(粗细曲线分别基于小波逼近和精确的方法). 由 δ 无偏估计量的渐进 Cramér-Rao 下界得到 $N=512$ 时 0.0344 的平方误差均值的平方根，通过从给定的估计值中减和加 0.0689 我们形成 δ 的大约 95%的置信区间(对极大似然估计变化的评估也告诉我们$\tilde{\delta}^{(s)}$与$\hat{\delta}$之间的偏差是无关紧要的).

图 387 展示了当 $j \geq 3$ 时通过式(344)而当 $j=1$ 和 2 时通过式(354)，$\tilde{\delta}^{(s)}$ 和 $\tilde{\sigma}_\varepsilon^2(\tilde{\delta}^{(s)})$可转变成 C_j 估计量(粗曲线)，另外还给出了基于$\widehat{\operatorname{var}}\{W_{j,t}\} \equiv \|\mathbf{W}_j\|^2/N_j$ 的 C_j 的样本估计量(圆圈). 对应于倍频程带中心频率的 C_j 的两个估计量都有绘出(与图 344 比较). 在所有的尺度上都有很好的一致性.

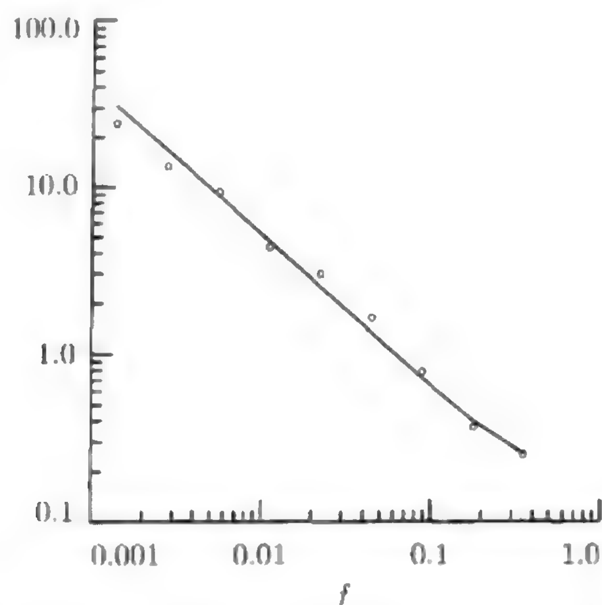


图 387　通过 LA(8)极大似然估计计算得到的小波系数方差$\tilde{\delta}^{(s)}$和$\hat{\sigma}_{\tau}^{2}(\tilde{\delta}^{(s)})$(实曲线)与 LA(8)小波系数样本方差(圆圈)做比较

表　387

τ_j	M'_j	D	临界水平		
			10%	5%	1%
1 年	331	0.155 9	0.094 5	0.105 1	0.126 2
2 年	165	0.175 4	0.132 0	0.146 9	0.176 5
4 年	82	0.100 0	0.185 5	0.206 8	0.247 4
8 年	41	0.231 3	0.257 2	0.286 4	0.343 6

注：对于方差齐次性使用哈尔小波滤波器，用计算机模拟确定临界值来检验尼罗河最低水位的结果(Whitcher 等，2000a).

现在，让我们转到 8.8 节讨论的使用实验统计 D 的方差齐次性假设估计问题. 为简单，使用哈尔小波(使用 D(4)和 LA(8)小波的结果实质上是同样的). 像在 9.6 节扩展与讨论[1]中所述，我们能够用同样的序列作这个实验，而不必是 2 的次幂($N=663$). 这里我们填上 1 024－663＝361 个零作为一个序列，之后在 $\mathbf{W}_j$ 中挑出前 $M'_j=\lfloor 663/2^j \rfloor$ 个小波系数(因为使用哈尔小波，M'_j 有一个简单公式). 对于前四个尺度的 D 统计值以及用计算机模拟的临界值在表 387 中给出. 因为对于 $\tau_1=1$ 年的 D 大于表中任一临界值，我们可以排斥在所有三个临界水平方差的齐次性零假设. 对于 $\tau_2=2$ 年，我们可以排斥水平 $\alpha=0.05$，但(勉强)不排斥水平 $\alpha=0.01$. 对于另外的尺度，我们不排斥任一重要的合理水平的零假设(这些结果与由图 327 给出的结论是一致的).

现在让我们转向研究用辅助统计量$\tilde{D}$来区别变动点的问题. 图 388 中间和下面的图表示在两个感兴趣的尺度上作为 k 的函数的 Δ_k^-，这(而不是 Δ_k^+)决定了$\tilde{D}$. 估计的变动点出现在很接近于 715 年新的测量仪器形成的时间. Beran and

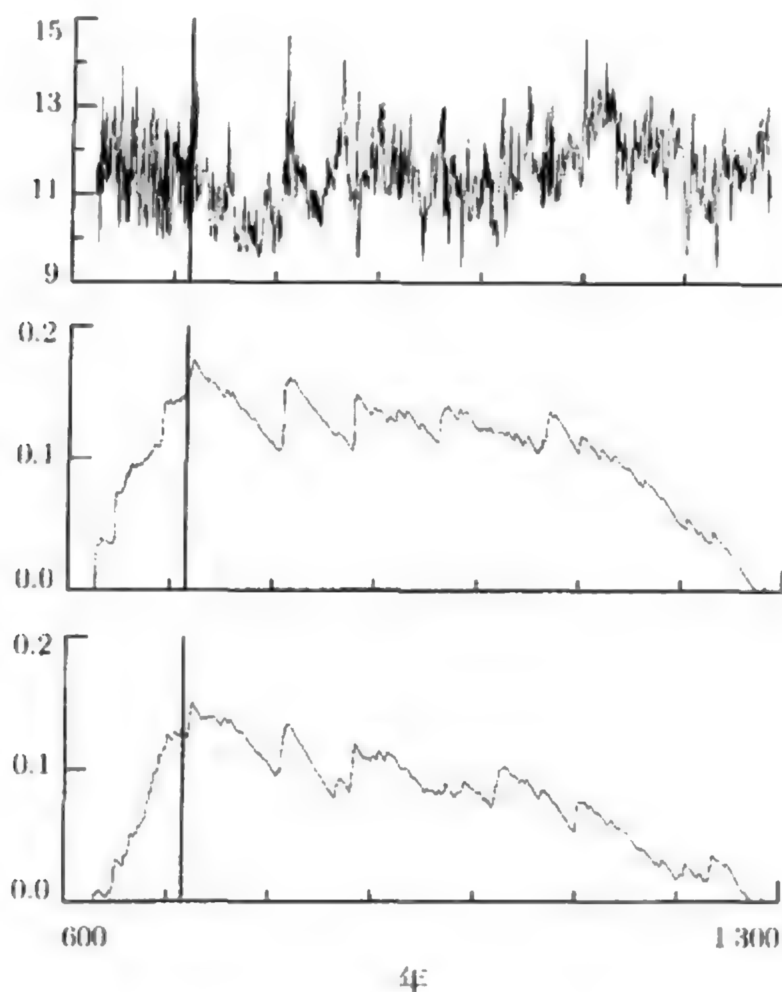


图 388 尼罗河最低水位(上面的图)以及对于尺度 τ_1 和 τ_2，作为 k 的函数的 Δ_k (中间和下面的图). 粗的临界线表示 715 年

Terrin(1996)也曾注意到尼罗河最低水位并且用了一个检验统计量讨论在时间序列中对长记忆参数的变动. 结合历史记录，这个分析的结果给出了一个备选解释，即减少是由于一个新的测量工具，而不是尼罗河长期性质的变化.

9.9 小结

令$\{X_t\}$是带有参数 δ 和 σ_ε^2 的零均值分形差分过程(精确的定义参见 7.6 节). 当 $0<\delta<\frac{1}{2}$ 时，这个过程是平稳的并且它的自协方差序列 $s_{X,\tau}\equiv\mathrm{cov}\{X_t, X_{t+\tau}\}$ 缓慢地趋于零的意义下，表现为长记忆，特别是对于接近 1/2 的 δ. 对于某个 $N=2^J$，令 $\mathbf{X}\equiv[X_0, X_1, \cdots, X_{N-1}]^T$，并且考虑 J 层(即完全的)离散小波变换 $\mathbf{W}\equiv\mathcal{W}\mathbf{X}$. 但是 $\mathbf{X}$ 的协方差矩阵 Σ_X 可能有显著的非零的非对角元素，而 $\mathbf{W}$ 的协方差矩阵 $\Sigma_W=\mathcal{W}\Sigma_X\mathcal{W}^T$ 的元素则相对较小. 因此我们可以主张离散小波变换对于分形差分和相关的平稳长记忆过程起到去相关变换的作用.

令 $C_j(j=1, \cdots, J)$是对于分形差分过程在与第 j 层小波系数相关的倍频程

带$\left[\frac{1}{2^{j+1}}, \frac{1}{2^j}\right]$上谱密度函数的平均水平：

$$C_j \equiv 2^{j+1}\int_{1/2^{j+1}}^{1/2^j} S_X(f)\mathrm{d}f.$$

对Σ_W的一个方便的逼近是对角矩阵Λ_N，它的对角线元素是

$$\underbrace{C_1,\cdots,C_1}_{\frac{N}{2}\uparrow},\underbrace{C_2,\cdots,C_2}_{\frac{N}{4}\uparrow},\cdots,\underbrace{C_j,\cdots,C_j}_{\frac{N}{2^j}\uparrow},\cdots,\underbrace{C_{J-1},C_{J-1}}_{2\uparrow},C_J,C_{J+1},$$

其中最后一个元素C_{J+1}由式(355b)给出，并且是对W中单个尺度系数方差的逼近．当X是多元高斯过程时，我们可以说它的离散小波变换W也是一个多元的高斯过程，并且有均值向量$\mathbf{0}$和能被Λ_N很好地逼近的方差．

W统计性质的这个逼近的两个用处是，综合具有长记忆的平稳分形差分过程和估计它们的参数．为了从分形差分过程中综合(即模拟)一个长度为$N/4$的时间序列，首先，用相应于想要的参数δ和σ_ϵ^2的谱水平C_j的平方根来装填它形成对角矩阵$\Lambda_N^{1/2}$；即有$\Lambda_N^{1/2}\Lambda_N^{1/2}=\Lambda_N$．其次，取具有零均值和单位方差的高斯白噪声过程的$N$个偏离构造向量$Z_N$，并建立

$$Y_N \equiv \mathcal{W}^{\mathrm{T}}\Lambda_N^{1/2}Z_N.$$

然后我们得到随机变量的实现κ，它是在整数$0,\cdots,N-1$上均匀分布的．我们使用这个来实现循环平移Y_N；即构造$\widetilde{Y}_N\equiv\mathcal{T}^{-\kappa}Y_N$，其中$\mathcal{T}$是循环平移矩阵(参见围绕式(52a)展开的讨论)．$\widetilde{Y}_N$的前$N/4$个元素是平稳过程的实现，这个平稳过程的自协方差序列与所期望的分形差分过程在延迟$\tau=0,\cdots,\frac{N}{4}-1$时非常匹配．

假设现在有一个时间序列X，我们想要用它来建立具有零均值的分形差分过程的模型．我们可以用W的近似统计性质来给出δ和σ_ϵ^2的简单极小似然估计量公式．给定X，这两个参数的精确的似然函数由

$$L(\delta,\sigma_\epsilon^2 \mid X) \equiv \frac{1}{(2\pi)^{N/2}\mid\Sigma_X\mid^{1/2}}\mathrm{e}^{-X^{\mathrm{T}}\Sigma_X^{-1}X/2}$$

给出．近似关系式$\Sigma_W=\mathcal{W}\,\Sigma_X\,\mathcal{W}^{\mathrm{T}}\approx\Lambda_N$可以使我们简化上面公式的两个主要部分．首先，我们有$\mid\Sigma_X\mid\approx\mid\Lambda_N\mid$，这转而等于$\Lambda_N$所有对角元素的乘积．其次，我们有$\Sigma_X^{-1}\approx\mathcal{W}^{\mathrm{T}}\Lambda_N^{-1}\mathcal{W}$．用这些替换，我们就导出对$\delta$的一个简化对数似然函数，即

$$\tilde{l}(\delta \mid X) \equiv N\log(\bar{\sigma}_\epsilon^2(\delta)) + \log(C'_{J+1}(\delta)) + \sum_{j=1}^{J} N_j\log(C'_j(\delta)),$$

其中

$$C'_j(\delta) \equiv 2^{j+1}\int_{1/2^{j+1}}^{1/2^j}\frac{1}{[4\sin^2(\pi f)]^\delta}\mathrm{d}f,\quad j=1,\cdots,J$$

(剩余变量 C_{j+1} 由式(362d)给出). δ 的极大似然估计是使得$\tilde{l}(\delta \mid \boldsymbol{X})$最小的$\tilde{\delta}^{(s)}$，而相应的 σ_ε^2 的极大似然估计则将$\tilde{\delta}^{(s)}$代入式(363b)的右边就能得到. 基于小波的近似极大似然估计$\tilde{\delta}^{(s)}$与精确的$\hat{\delta}$相比十分有利(参见表 363 和表 364). 就像练习[365b]指出的那样，我们可以很容易地采用这种方法来处理预先没有假定有零均值的时间序列.

对给定的时间序列 $\boldsymbol{X}$，极大似然估计$\tilde{\delta}^{(s)}$是基于公式化逼近对 δ 的似然函数来给出的；然而，对给定 $\boldsymbol{X}$ 的离散小波变换 $\boldsymbol{W}$，我们仍然可以解释 δ 的近似似然函数并以此作为结束. 如果限制仅用 $\boldsymbol{W}$ 中的非边界小波系数，我们可以构造 δ 的似然函数，这个 δ 可以适合(i)非平稳的分形差分过程(ii)在确定的多项式趋势之前观察到的平稳或非平稳分形差分过程. 将相应的简化对数似然函数极小化就能得到基于小波的近似极大似然估计$\tilde{\delta}^{(s/ns)}$. 在精确的极大似然估计$\hat{\delta}$和近似的极大似然估计$\tilde{\delta}^{(s)}$可以应用的情况下(即无趋势的平稳分形差分过程)，这个估计量有比它们大得多的平方误差均值的平方根，但是$\tilde{\delta}^{(s/ns)}$可以直接应用到比另外两个极大似然估计更宽的方法类中. 例如，在模型 $\boldsymbol{X}=\boldsymbol{T}+\boldsymbol{U}$ 中，我们可以在检验 $\boldsymbol{T}=\boldsymbol{0}$ 的假设时用$\tilde{\delta}^{(s/ns)}$，其中 $\boldsymbol{T}$ 是一个多项式趋势，而 $\boldsymbol{U}$ 是一个零均值分形差分过程. 检验基于 $\boldsymbol{W}$ 中尺度和边界小波系数形成趋势估计$\hat{\boldsymbol{T}}$，然后计算检验统计量 $P\equiv\|\boldsymbol{X}\|^2/\|\boldsymbol{X}-\hat{\boldsymbol{T}}\|^2$. 以观察值$\tilde{\delta}^{(s/ns)}$为条件，我们根据经验可以通过模拟研究确定零假设下 P 的分布，然后评估是接受还是拒绝零假设.

作为极大似然估计的一种选择，我们也可以用最小平方估计量来估计 δ. 这个思想是基于对分形差分过程来说作为尺度 τ_j 函数的小波方差 $\nu_X^2(\tau_j)$的 log/log 图是近似线性的这一事实的. 这个近似在 $\tau_j\to\infty$时变得更加精确，在除 τ_1 以外的所有尺度上都能很好地接受. 因此对于 $j\geqslant 2$，我们可以写 $\log(\nu_X^2(\tau_j))\approx\zeta+\beta\log(\tau_j)$，其中斜率 β 和 δ 有关系 $\delta=\frac{1}{2}(\beta+1)$. 给定一个长度为 L 的小波滤波器和一个可以视为有 $\delta\in\left[-\frac{1}{4},\ \frac{L}{2}+\frac{1}{4}\right]$的分形差分过程实现的时间序列(参见围绕式(370b)的讨论)，我们用 M_j 个非边界极大重叠离散小波变换小波系数形成小波方差的无偏估计量$\hat{\nu}_X^2(\tau_j)$(参见式(306b)). 式(313a)的分布结果记作

$$\hat{\nu}_X^2(\tau_j)\overset{\mathrm{d}}{=}\nu_X^2(\tau_j)\chi_{\eta_j}^2/\eta_j,$$

其中，为了简便，近似地取等价自由度 η_j 作为 $\max\{M_j/2^j,\ 1\}$(这与式(314c)一致). 现在我们可以给出线性回归模型的公式

$$Y(\tau_j)=\zeta+\beta\log(\tau_j)+e_j,$$

其中

$$Y(\tau_j)\equiv\log(\hat{\nu}_X^2(\tau_j))-\psi\left(\frac{\eta_j}{2}\right)+\log\left(\frac{\eta_j}{2}\right),$$

和 $\psi(\cdot)$是 Γ_2 函数. 在这里，误差项 e_j 近似地认为是不相关的，并且与随机变量 $\log(\chi_{\eta_j}^2)-\psi\left(\frac{\eta_j}{2}\right)-\log(2)$有相同的分布；进而，$e_j$ 有零均值和方差 $\psi'\left(\frac{\eta_j}{2}\right)$，并且当

等价自由度在 10 附近或大于 10 时近似于高斯分布，其中 $\psi'(\cdot)$是 Γ_3 函数. 标准最小二乘理论使我们得到式(376c)中给出的加权最小二乘估计$\hat{\beta}^{(\text{wls})}$(对于加权最小二乘估计理论上的方差由式(377)给出). 相应的 δ 估计量由$\hat{\delta}^{(\text{wls})}\equiv\frac{1}{2}(\hat{\beta}^{(\text{wls})}+1)$给出. 从表 377 可以看出，加权最小二乘估计有比精确的极大似然估计大得多的平方误差均值的平方根，前者很容易计算并且——与基于小波的近似极大似然估计$\tilde{\delta}^{(\text{s/ns})}$类似——对(i)非平稳的分形差分过程(ii)在确定的多项式趋势出现的时候观察到的平稳或非平稳分形差分过程有直接应用的优势.

离散小波变换近似去相关分形差分过程的事实也可以为具有长记忆结构的时间序列的方差齐次性设计一个检验. 基本思想是，在 $\mathbf{X}$ 是高斯分形差分过程的一个样本这样的零假设下，$\mathbf{W}_j$ 中的非边界小波系数应当近似地是一个具有齐次方差的高斯白噪声过程的一个样本. 利用通过式(380b)定义的检验统计量 D，我们可以在每个尺度对尺度的基础上检验齐次方差的零假设. 如果我们对第 j 层小波系数拒绝零假设，并且接受时间序列本身方差的一个单个变动替换的假设，那么就可以用一个基于极大重叠离散小波变换的小波系数的辅助统计量$\widetilde{D}$来估计变动点的位置(参见式(382)及围绕它展开的讨论).

9.10 练习

[9.1] 类似于图 347a 和 347b，对 D(4)离散小波变换绘图.

[9.2] 证明，如果我们将$\mathcal{H}_j(\cdot)$近似地看作是一个理想带通滤波器的平方增益函数，并且如果假设 $S_X(\cdot)$在这个通带上是常数，则从式(348b)得出 $\operatorname{cov}\{W_{j,t}, W_{j,t+\tau}\}\approx 0$.

[9.3] 令$\{X_t\}$是一阶移动平均(MA)过程；即我们能够写

$$X_t=\varepsilon_t+\theta\varepsilon_{t-1},$$

其中$\{\varepsilon_t\}$是白噪声过程，θ是一个任意的常数. 得到这个过程的自协方差序列和谱密度函数，然后类似于图 349 那样，对 $\theta=0.5$ 和 $\theta=0.98$ 的情形绘图. 离散小波变换如何去相关这些移动平均过程？(对于$-1<\theta<1$ 的移动平均过程的讨论参见 Whitcher，1998，4.1.2 节.)

[9.4] 假定随机变量的 N 维向量 $\mathbf{Y}$ 服从具有所有分量全为零的均值向量和协方差矩阵 Σ_Y 的多变量高斯分布，这个矩阵的第(m, n)个元素用 $\Sigma_{Y,m,n}$表示. 令 κ 是均匀分布在0，1，…，$N-1$上的一个随机变量，并且考虑$\widetilde{\mathbf{Y}}\equiv\mathcal{T}^{-\kappa}\mathbf{Y}$，其中$\mathcal{T}$是式(52a)的循环位移矩阵(回忆$\mathcal{T}^2=\mathcal{T}\mathcal{T}$等等). 证明：$\widetilde{\mathbf{Y}}$中的随机变量构成一个平稳过程的一段，它的自协方差序列由式(357)给出. 再证明：已知对于 $\tau=1$，…，$N/2$，在 $s_{\widetilde{Y},N-\tau}=s_{\widetilde{Y},\tau}$的意义上，这个自协方差序列关于 $\tau=N/2$ 对称.

[9.5] 验证式(376c)和(377).

[9.6] 验证式(374)和(379)，它们分别可以用来得到 δ 的瞬时极大似然和最小二乘估计量. 并确定在 LA(8)小波情况下对于 t_j 的表示. 最后，讨论 δ

的瞬时最小二乘估计值的方差的表达式是

$$\frac{(J_0-J_1+1)\pi^2}{8[(J_0-J_1+1)\sum\log^2(\tau_j)-(\sum\log(\tau_j))^2]}.$$

[9.7] 绘类似于图383的图，但是现在使用D(4)和C(6)小波. 将估计趋势$\hat{\boldsymbol{T}}$与图383中的趋势做比较，然后讨论每个小波对得到的趋势估计有什么影响.

[9.8] 基于尼罗河最低水位的最后$N=512$个数据，使用哈尔和D(4)小波计算一个平稳分形差分过程参数δ和σ_ε^2的基于小波的近似极大似然估计. 用这些极大似然估计绘类似于图387的图. 哈尔和D(4)极大似然估计与9.8节给出的LA(8)极大似然估计相比各自的优缺点是什么?（尼罗河序列可以通过这本书的网页得到）

第10章

基于小波的信号估计

10.0 引言

在第4章中已经介绍的离散小波变换可以把一个时间序列 $\mathbf{X}$ 分析(分解)为离散小波系数 $\mathbf{W}$，同时我们可以从 $\mathbf{W}$ 综合(重构)原始序列. 我们已经注意到，例如综合相位可以用来重构时间序列的多分辨分析(参见式(64)或者(104a))和模拟长记忆过程(参见9.2节). 本章我们要研究对于综合相位的另一个重要的用法：适用于信号估计(或者称为函数估计或去噪)问题，这里我们想在观察时间序列中估计一个被噪声包含的信号. 这里，基本的想法就是修改 $\mathbf{W}$ 中的元素以产生 $\mathbf{W}'$，从而信号的估计可以综合出来. 除了在10.8节简单介绍的方法以外，一旦某个参数已经估计出来，$\mathbf{W}$ 的元素 W_n 就一次一个地被处理；即我们怎样修改 W_n 并不受保留下来的离散小波变换系数的直接影响. 我们这里讲的基于小波的技术在概念上都是非常简单的，然而却适合对广泛的各式各样的信号进行处理.

我们将考虑确定性信号和统计信号，也考虑不相关噪声和相关噪声. 我们采用下面的符号约定来强调要讨论的模型：

[1]$\mathbf{D}$ 表示一个 N 维确定信号向量.

[2]$\mathbf{C}$ 表示一个 N 维统计(随机)信号向量，即 $\mathbf{C}$ 是一个多元随机变量，协方差矩阵为 Σ_C.

[3]$\boldsymbol{\varepsilon}$ 表示 N 维独立同分布(IID)均值为零的噪声向量；即 $\boldsymbol{\varepsilon}$ 是一个多元随机变量，每个元素有相同的不变分布，$E\{\boldsymbol{\varepsilon}\}=\mathbf{0}$(一个 N 维零向量)，协方差矩阵 $\Sigma_\varepsilon=\sigma_\varepsilon^2 I_N$，其中 σ_ε^2 是普通方差.

[4]$\boldsymbol{\eta}$ 表示 N 维非独立同分布的噪声向量，均值为零. 即 $\boldsymbol{\eta}$ 是一个多元随机变量且 $E\{\boldsymbol{\eta}\}=\mathbf{0}$，并且协方差矩阵是 Σ_η(依赖于文中的特别假设，$\boldsymbol{\eta}$ 中的随机变量可以从具有不同方差的独立同分布或者是相关的分布中产生).

本章我们将在适当的位置提出有关随机变量性质的假设.

本章组织如下. 首先我们将考虑应用于确定性信号的一般的规范正交变换 $\mathcal{O}$. 在 10.1 节，我们将指出实际上一般规范正交变换简洁地概括一个信号所需要的系数比其他的更少(特别地，多样的确定信号 $\boldsymbol{D}$ 的离散小波变换可以将 $\boldsymbol{D}$ 的主要性质集中到少量的系数上). 这种推测导致了的想法——在 10.2 节中所展示——对附加噪声的信号的规范正交变换采用阈值方法. 我们讨论了不同类型的阈值函数，同时阐述了选择阈值层的技术. 在 10.3 节中我们说明在噪声下统计信号 $\boldsymbol{C}$ 的最小均方估计可以通过一个对观察系数的简单尺度得到，从而得到一个估计(与著名的维纳(Wiener)滤波器相似)，即将观察到的变换系数收缩到零. 我们讨论的收缩法则完全展示了在 10.4 节中信号变换系数和噪声变换系数的统计结构——这些包括条件平均、中值方法和贝叶斯(Bayes)公式. 我们对观察到的离散小波变换系数应用阈值或者收缩方法所得到的严格依赖于噪声变换系数的结构. 当时间序列中的附加的噪声项是独立同分布的高斯分布时，噪声变换系数也是，我们将在 10.5 节验证这种情况. 而 10.6 节我们观察附加噪声项是不相关、非高斯的情况，这样噪声系数的离散小波变换不是独立同分布的，从而阈值成了问题，对这种结构一个有趣的应用就是用周期化进行谱估计. 在 10.7 节我们将考虑在一种很特别的例子中的相关高斯噪声系数的离散小波变换，这个模型是个很好的逼近，称为多锥谱估计基础上的谱估计. 在 10.5 节到 10.7 节中集中于处理的噪声过程 ϵ 或 η 的结构，我们将在 10.8 节讨论对于统计信号 $\boldsymbol{C}$ 的一个逼近. 在 10.9 节我们做了一个小结.

10.1 信号的小波表示

$\boldsymbol{D}$ 为 N 维向量，其元素组成感兴趣的确定性信号. 本节的中心议题是 $\boldsymbol{D}$ 的离散小波变换能够表示(重新表示)信号的有效方法. 这种方式易于简化信号估计的问题. 如我们将在 10.2 节看到的，就是能够使信号估计的表示用很少数目，即 M 项，本质地总结出 $\boldsymbol{D}$ 的重要性质. 如果 M 比 N 小得多，那么估计问题就简化为只需估计 $\boldsymbol{D}$ 中的 M 项，而不是所有 N 项. 我们所寻找的就是用很少的项来表示信号 $\boldsymbol{D}$ 的有效方法，而我们要求的这些离散小波变换就可以达到.

为了检验相对于其他的规范正交变换，离散小波变换怎样更好地表示信号 $\boldsymbol{D}$，让我们做一个平凡假设 $\|\boldsymbol{D}\|^2>0$(即排除信号恒等于零的情况). 我们能运用符号规范的部分能量序列(NPES)来看看，一个规范正交变换是如何以少量的项来获取 $\boldsymbol{D}$ 的主要性质的. 如前面 4.10 节中描述的，如果 $\boldsymbol{O}=\mathcal{O}\boldsymbol{D}$ 是由信号 $\boldsymbol{D}$ 的一个实值或者复值的规范正交变换 $\mathcal{O}$ 形成的变换系数的向量，而 $\boldsymbol{O}$ 的规范部分能

量序列实如下定义. 令 $O_{(l)}$ 是 $\mathbf{O}$ 中元素 $O_{l'}$，第 l 个元素即那些平方量 $|O_{l'}|^2$ 是在 N 个可能的平方量中是第 l 大的. 因此我们有

$$|O_{(0)}|^2 \geqslant |O_{(1)}|^2 \geqslant \cdots \geqslant |O_{(N-1)}|^2.$$

下面我们对项 $|O_{(l)}|^2$ 递增求和，并用能量 $\|\mathbf{O}\|^2 = \|\mathbf{D}\|^2$ 规范化这个和来定义规范的部分能力序列 $\{C_n\}$ 如下：

$$C_n \equiv \frac{\sum_{l=0}^{n} |O_{(l)}|^2}{\|\mathbf{O}\|^2} = \frac{\sum_{l=0}^{n} |O_{(l)}|^2}{\sum_{l=0}^{N-1} |O_{(l)}|^2}, \quad n = 0, \cdots, N-1.$$

由构造，$0 < C_n \leqslant 1$ 且 $C_n \leqslant C_{n+1}$. 为了看一下 $\mathbf{O}$ 中具有最大平方量的 M 个系数是怎样提取 $\mathbf{D}$ 中的性质，令 $\mathcal{I}_M$ 是 $N \times N$ 阶对角矩阵，其元素由 M 个 1(指 M 个最大 $|O_{l'}|^2$ 的位置)和 $N-M$ 个零组成. 令 $\hat{\mathbf{D}}_M \equiv \mathcal{O}^{\mathrm{T}} \mathcal{I}_M \mathbf{O}$；即 $\hat{\mathbf{D}}_M$ 是由 M 个最大的变换系数形成的 $\mathbf{D}$ 的一个逼近.

练习[395] 证明

$$C_{M-1} = 1 - \frac{\|\mathbf{D} - \hat{\mathbf{D}}_M\|^2}{\|\mathbf{D}\|^2}. \qquad \triangleleft$$

注意 $1 - C_{M-1}$ 表示当我们用 $\hat{\mathbf{D}}_M$ 逼近 $\mathbf{D}$ 的时候的相对逼近误差，即逼近误差量的平方和与信号量平方和的比值. 因此，$1 - C_{M-1}$ 越小(或者等价地 C_{M-1} 越大)逼近就越好. 如果 $\{C_n\}$ 和 $\{C'_n\}$ 是两组规范正交变换的规范的部分能量序列，而且如果 $C_{M-1} > C'_{M-1}$，在恰当运用 M 个变换系数来表示算法的时候，我们将更倾向于第一个变换而不是第二个变换.

为了看一下为什么离散小波变换对于信号表示是一个很好的规范正交变换，让我们考虑图 396中左列的三个信号. 每一个信号 $\mathbf{D}_j$ 都是由 $N = 128$ 个元素组成，表示为 $D_{j,t}$，$t = 0$，…，127. 从上到下，三个信号依次为

[1]一个取样的正弦曲线 $D_{1,t} = \frac{1}{2}\cos\left(\frac{3\pi t}{32} + 0.08\right)$.

[2]一个“脉冲”对所有的 t 定义为 $D_{2,t} = 0$，除去 $t = 59$，…，69 的点，在这些点上 $D_{2,t} = D_{1,t}$.

[3]$D_{3,t} = \frac{1}{\sqrt{120}} D_{1,t} + D_{2,t}$.

定性地讲，$\mathbf{D}_1$ 是一个“频域”信号(因为它的规范正交离散傅里叶变换中仅由两个非零的系数组成)，而 $\mathbf{D}_2$ 是“时域”信号(因为它仅有 11 个非零值). 第三个信号 $\mathbf{D}_3$ 是这两个域的混合. 右列的图形展示的是，每个信号分别用了三种不同的规范正交变换的规范的部分能量序列，这三个变换是恒等变换 I_N(点划曲线)，正交离散傅里叶变换 $\mathcal{F}$(细实线)和基于 LA(8)Daubechies 小波滤波器的离散小波变换 $\mathcal{W}$(虚线). 表 396 表示 M 的最小值，产生了一个对信号和变换的各种复合不超过 1%的相对逼近误差. 而离散小波变换在适用于单位变换和正交离散傅

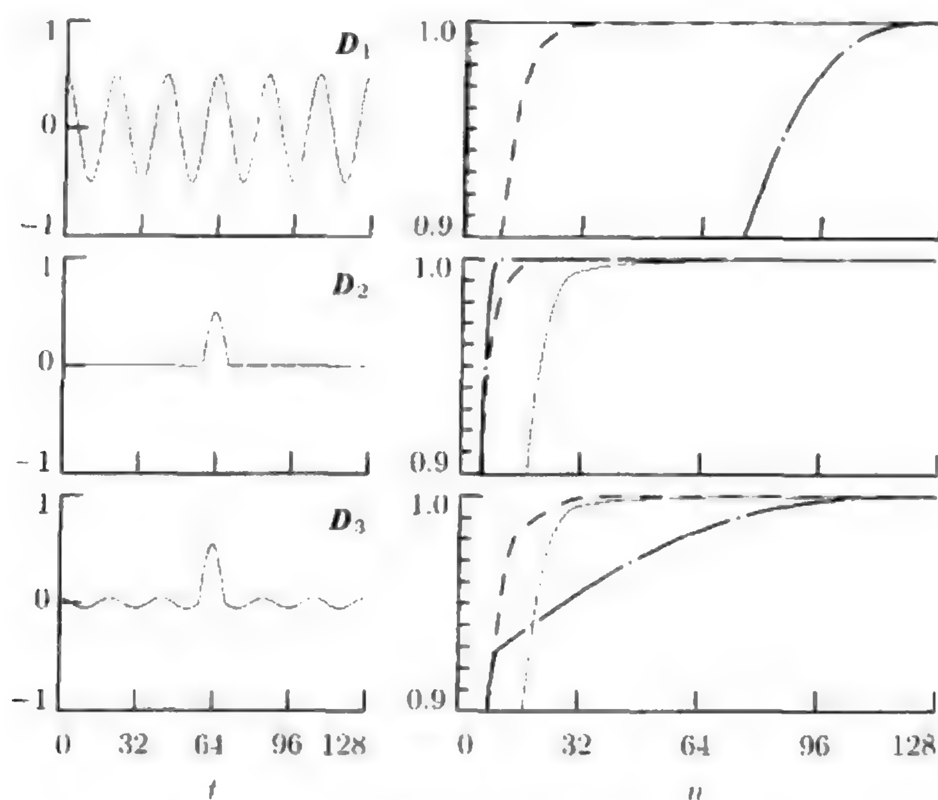


图 396 三个信号(左列)和它们相应的规范化部分能量序列(右列)对信号本身(点划线)，它们的正交离散傅里叶变换系数(细实线)和 LA(8)离散小波变换系数(虚线)

表 396

	信号域		
	频域	时域	混合
$\mathcal{F}$	2	29	28
I_N	105	9	75
$\mathcal{W}$	22	14	21

注：对三个信号用正交离散傅里叶变换$\mathcal{F}$、恒等变换 I_N 和 LA(8)离散小波变换$\mathcal{W}$，获得不超过 1% 的相对逼近误差所需要的系数的数目(详情参见正文)。

里叶变换的信号时次于这两种变换，离散小波变换在两种情况中都是第二的，而它对混合信号实际上是相当好的．这个简单的例子表明离散小波变换应当对实际发生的很多信号即那种表现为时域特征(短暂的)和频域特征(窄带和宽带特征)的混合信号起到很好的效果．

作为一个离散小波变换如何很好地表示实际相当复杂的信号的例子，让我们考虑海洋测量的垂直截面的一部分，如图 194 的最下面的和在 5.10 节中描述的(这个数据实际上不能认为是一个确定信号，但是我们在这里无论如何要这么做以描述关于离散小波变换的要点)．为了使用 $J_0=7$ 层部分 LA(8)离散小波变换，我们需要一个为 $2^{J_0}=128$ 的倍数的样本尺寸，因此我们缩减原来序列 91 个观察

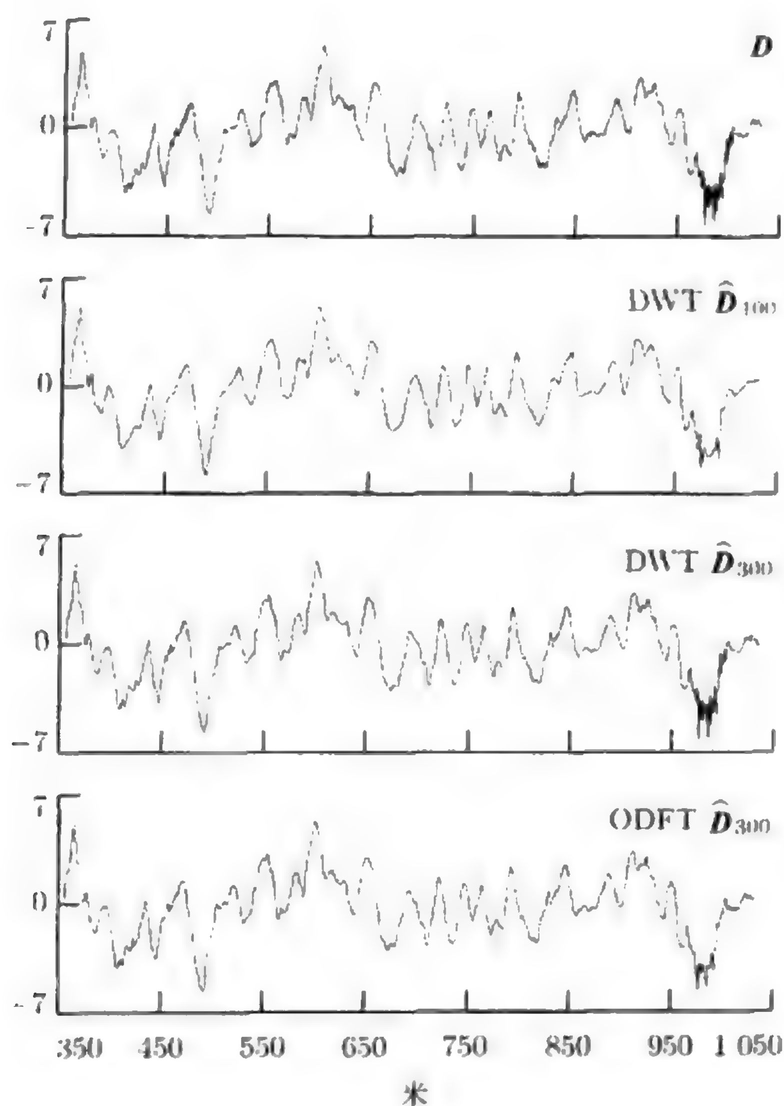


图 397　绘出了与对应于深度为米的海洋垂直截面的假设信号 $\mathbf{D}$ 的 $N=6\ 784$ 个值(顶部的图)，连同用 100 个 LA(8)离散小波系数、300 个 LA(8)离散小波系数和 300 个正交离散傅里叶变换系数组成的重构. 对这些数据的更多讨论参见 5.10 节

点来获取子序列长度为 $N=6\ 784=53\cdot 128$——这个子序列在图 397 的最上面给出(如果我们将原序列表示为 X_0，…，$X_{6\ 871}$，那么这个子序列为 X_{54}，…，$X_{6\ 837}$). 在这个图下面的三个图表示了 $\hat{\mathbf{D}}_M$，分别基于 100 个 LA(8)离散小波变换系数、300 个 LA(8)离散小波变换的系数和 300 个正交离散傅里叶变换系数. 直觉上认为保留短暂事件用 300 个离散小波变换系数逼近要优于 300 个正交离散傅里叶变换(注意 300 个系数比 N 的 5%还少). 借助于规范部分能量序列，我们有 $C_{299}^{(\mathrm{dwt})}=0.998\ 3$ 和 $C_{299}^{(\mathrm{odft})}=0.997\ 3$，因此，在这种度量下的离散小波变换优于正交离散傅里叶变换(事实上我们有 $C_{122}^{(\mathrm{odft})}<C_{299}^{(\mathrm{dwt})}<C_{123}^{(\mathrm{odft})}$，因此我们需要另外 123 个正交离散傅里叶变换系数以获得与 LA(8)$\hat{\mathbf{D}}_{300}$ 相同的逼近误差).

10.2 通过阈值方法的信号估计

现在假设我们把时间序列模型化为 $\boldsymbol{X}=\boldsymbol{D}+\varepsilon$，其中 $\boldsymbol{D}$ 表示一个感兴趣的未知的确定性信号，而 ε 是不想要的独立同分布均值为零的统计噪声．再次，我们令$\mathcal{O}$为 $N\times N$ 阶规范正交变换矩阵，我们取 $\boldsymbol{O}\equiv\mathcal{O}\boldsymbol{X}$ 来表示变换系数 $\{O_l: L=0,\cdots,N-1\}$的 N 维向量．那么

$$\boldsymbol{O}\equiv\mathcal{O}\boldsymbol{X}=\mathcal{O}\boldsymbol{D}+\mathcal{O}\varepsilon\equiv\boldsymbol{d}+\boldsymbol{e}\quad \text{因此 } O_l=d_l+e_l, \tag{398}$$

其中 d_l 和 e_l 分别是 $\boldsymbol{d}=\mathcal{O}\boldsymbol{D}$ 和 $\boldsymbol{e}=\mathcal{O}\varepsilon$ 的第 l 个元素．我们能够定义这种模型的信噪比为信号范数的平方和噪声范数平方的期望的比值：$\|\boldsymbol{D}\|^2/E\{\|\varepsilon\|^2\}=\|\boldsymbol{d}\|^2/E\{\|\boldsymbol{e}\|^2\}$．如果这个比率大并且变换$\mathcal{O}$能成功地孤立信号使得信号变换系数 $\boldsymbol{d}$ 的一小部分是大的，那么 $\boldsymbol{O}$ 应当由少数大系数组成，这些大系数对信号起决定作用，而许多小系数来自于噪声．

假设 M 是归因于信号的未知系数的数目，$\mathcal{I}_M$ 是除掉一些来自 $\boldsymbol{O}$ 的系数的 $N\times N$ 阶矩阵，那么我们的问题就是如何确定 M，还有，给出 M 来构造 $\mathcal{I}_M$，由这些我们可以用

$$\hat{\boldsymbol{D}}_M\equiv\mathcal{O}^{\mathrm{T}}\mathcal{I}_M\boldsymbol{O}$$

来估计信号．这个问题的解可以通过最小化标准

$$\gamma_m\equiv\|\boldsymbol{X}-\hat{\boldsymbol{D}}_m\|^2+m\delta^2,\quad m=0,\cdots,N$$

来得到．给定 m，通过所有可能的$\mathcal{I}_m$，其中 $\delta^2>0$ 是常数且可以选择(这里我们定义$\hat{\boldsymbol{D}}_0=\boldsymbol{0}$)．在 γ_m 中的第一项能认为是逼真度条件，这就是说我们想要估计的 $\boldsymbol{D}$ 离观察数据不太远．注意到当 $m=N$ 时这个逼真度条件是零而且在 $m=0$ 时获得最大值$\|\boldsymbol{X}\|^2$．如果对所有的 m 有 $\mathcal{I}_{m+1}-\mathcal{I}_m=\mathcal{I}_1$，这个逼近条件就成为 m 的非增函数．γ_m的第二项是阈值项，减轻在构造$\hat{\boldsymbol{D}}_m$中使用的大量的系数，是 m 的严格递增函数．

我们断定 γ_m 在 $m=M$ 时最小，如果我们确定 M 是使$|O_l|^2>\delta^2$变换系数的数目．为此，令$\mathcal{J}_m$为 m 指标 l 的集合，l 是$\mathcal{I}_m$的第 l 个对角元素等于 1 的指标，并且注意到

$$\begin{aligned}\gamma_m&\equiv\|\boldsymbol{X}-\hat{\boldsymbol{D}}_m\|^2+m\delta^2=\|\mathcal{O}^{\mathrm{T}}\boldsymbol{O}-\mathcal{O}^{\mathrm{T}}\mathcal{I}_m\boldsymbol{O}\|^2+m\delta^2\\&=\|(I_N-\mathcal{I}_m)\boldsymbol{O}\|^2+m\delta^2=\sum_{l\notin\mathcal{J}_m}|O_l|^2+\sum_{l\in\mathcal{J}_m}\delta^2.\end{aligned}$$

注意到，如果 $l\in\mathcal{J}_m$，我们贡献 δ^2到组成 γ_m 的和式里，然而，如果 $l\notin\mathcal{J}_m$，我们贡献$|O_l|^2$到和式里面．因此我们通过把所有满足$|O_l|^2>\delta^2$的 l 放到$\mathcal{J}_m$中最小化 γ_m．这个结果表明，最小化标准 γ_m 通过 m 自然地导致了基于变换系数的阈值的一个估计$\hat{\boldsymbol{D}}_M$(Moulin，1995)．

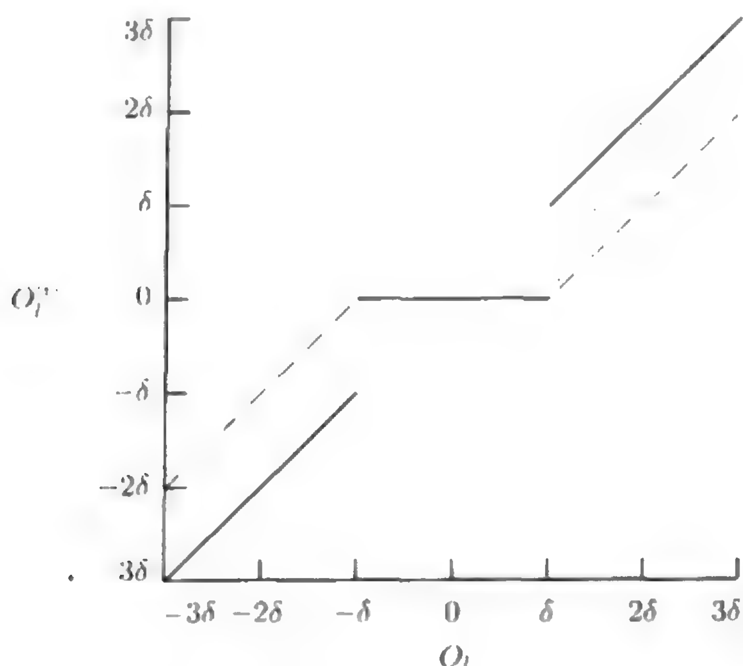


图 399 映射 O_l 到 $O_l^{(t)}$，其中 $O_l^{(t)}$ 是 $O_l^{(ht)}$ 对硬阈值(实线)，或者是 $O_l^{(st)}$ 对软阈值(虚线)，或者是 $O_l^{(mt)}$ 对中阈值(点线). 注意所有三个阈值算法当 $-\delta\leqslant O_l\leqslant\delta$ 时候是一样的

• 阈值函数

在阈值函数公式化中的一个普通线索是，如果观察的变换系数 O_l 的幅值小于阈值 $\delta>0$，系数就置为零. 因此对于估计 $\boldsymbol{D}$ 的一个阈值算法是由下面三个基本步骤组成的.

[1]计算变换系数 $\boldsymbol{O}\equiv\mathcal{O}\boldsymbol{X}$.

[2]定义阈值系数 $\boldsymbol{O}^{(t)}$ 是一个向量，它的第 l 个元素是 $O_l^{(t)}$.

$$O_l^{(t)}=\begin{cases}0, & |O_l|\leqslant\delta;\\ \text{某些非零值}, & \text{其他}.\end{cases}$$

这里的非零值还没有定义.

[3]通过 $\hat{\boldsymbol{D}}^{(t)}\equiv\mathcal{O}^{\mathrm{T}}\boldsymbol{O}^{(t)}$ 估计 $\boldsymbol{D}$.

注意同样的阈值应用到每个项 O_l. 这个类型的阈值有时称为全局阈值.

对于幅值超过 δ 的系数，有几种可能性. 最简单的是硬阈值，我们定义系数 $\boldsymbol{O}^{(t)}$ 阈值后为 $\boldsymbol{O}^{(ht)}$，其元素是

$$O_l^{(ht)}=\begin{cases}0, & |O_l|\leqslant\delta;\\ O_l, & \text{其他}.\end{cases}\tag{399}$$

这里幅值超过 δ 的系数保持不变，而那些小于或者等于 δ 的被“杀”或设置为零. 从 O_l 到 $O_l^{(ht)}$ 的映射用图 399 中的实线表示.

另一种普遍应用的阈值是软阈值，这里 $\mathbf{O}^{(t)}$ 等于向量 $\mathbf{O}^{(st)}$，其第 l 个元素为

$$O_l^{(st)} = \mathrm{sign}\{O_l\}(|O_l| - \delta)_+, \tag{400a}$$

其中

$$\mathrm{sign}\{O_l\} \equiv \begin{cases} +1, & O_l > 0; \\ 0, & O_l = 0; \\ -1, & O_l < 0; \end{cases}$$

而且

$$(x)_+ \equiv \begin{cases} x, & x \geqslant 0; \\ 0, & x < 0. \end{cases}$$

系数的数值超过 δ 则以 δ 推向零，而那些小于或者等于 δ 则置为零．图 399 中的虚线描绘了从 O_l 到 $O_l^{(st)}$ 的映射．

在硬阈值和软阈值之间的一个妥协是中阈值，对于 $|O_l| \geqslant 2\delta$，跟硬阈值一样，对于 $\delta \leqslant |O_l| \leqslant 2\delta$，在硬阈值和软阈值之间做插——这个插值在图 399 中用点线表示出来了．这里我们设 $\mathbf{O}^{(t)}$ 等于向量 $\mathbf{O}^{(mt)}$，其元素为

$$O_l^{(mt)} = \mathrm{sign}\{O_l\}(|O_l| - \delta)_{++}, \tag{400b}$$

其中

$$(|O_l| - \delta)_{++} \equiv \begin{cases} 2(|O_l| - \delta)_+, & |O_l| < 2\delta; \\ |O_l|, & \text{其他}. \end{cases}$$

这里大系数，即那些数量超过 2δ 的系数保持不变，在 2δ 和 δ 之间的萎缩了，而小于 δ 的则置为零．

中阈值是更广的坚阈值(Bruce and Gao，1996b)的一种特殊情况，在对坚阈值步骤[2]要替换为由一个有两个阈值参数 δ 和 δ' 的阈值过程来定义的．这里阈值系数的向量由 $\mathbf{O}^{(ft)}$ 给出，其元素为

$$O_l^{(ft)} = \begin{cases} 0, & |O_l| \leqslant \delta; \\ \mathrm{sign}\{O_l\} \dfrac{\delta'(|O_l| - \delta)}{\delta' - \delta}, & \delta < |O_l| \leqslant \delta'; \\ O_l, & |O_l| > \delta'. \end{cases} \tag{400c}$$

因此我们集中看一下建立在阈值水平 δ 上的阈值法则．对于如何取 δ 的若干方法在文献中已经有了．

• 统一阈值

产生阈值水平的一个非常感兴趣的方法是由 Donoho and Johnstone(1994)对于独立同分布高斯噪声 ε 给出的，即 ε 是多元高斯(规范)随机变量，它的均值为零，方差为 $\sigma_\varepsilon^2 I_N$．关于独立同分布高斯噪声的规范正交变换的一个关键的性质就是变换后的噪声和未变换的噪声有相同的统计性质(参见练习[263])，这告诉我们在等式(398)中的 $\mathbf{e}$ 是一个独立同分布高斯随机变量且均值为零、方差为 σ_ε^2 的向量，即

$$e_l \stackrel{\mathrm{d}}{=} \mathcal{N}(0, \sigma_\varepsilon^2) \text{ 且当 } l \neq l' \text{ 时}, \mathrm{cov}\{e_l, e_{l'}\} = 0.$$

Donoho and Johnstone(1994)给出的 δ 的形式是

$$\delta^{(u)} \equiv \sqrt{[2\sigma_\varepsilon^2 \log(N)]}, \tag{400d}$$

这就是著名的统一阈值(上面的 log 是以 e 为底的). 这个阈值的基本原理如下. 假设信号 $\mathbf{D}$ 实际上是一个零向量，因此变换系数$\{O_l\}$是均值为零方差为 σ_ϵ^2 的独立同分布高斯序列$\{e_l\}$的一项，即 $\mathbf{O}=\mathbf{e}$. 那么当 $N\to\infty$时，我们有

$$\mathbf{P}[\max_l\{|O_l|\}\leqslant\delta^{(u)}]=\mathbf{P}[\max_l\{|e_l|\}\leqslant\delta^{(u)}]\to 1,$$

所以我们就渐进精确地估计信号向量. 这个结果可用两个理由小心地解释. 首先，阈值是倾向于除掉噪声的大部分，对于独立同分布高斯序列$\{e_l\}$，我们有

$$\mathbf{P}[\max_l\{|e_l|\}>\delta^{(u)}]\leqslant\frac{1}{\sqrt{[4\pi\log(N)]}},$$

因此在此范围中没有噪声可以通过这个阈值. 对于有限采样，这个概率界限说明不超过 13%的长度为 $N=128$ 或更大的纯噪声变量的实现将超过这个阈值(Johnstone and Silverman，1997，p. 325). 因此统一阈值可以移去所有的噪声. 但是，在这样做的时候它也可能无意地使得某些小信号的变换系数为零. 因为这样倾向于光滑信号，因此统一阈值以极大的概率保证了重构后的信号和真正的确定信号一样光滑或者更光滑.

第二，如果 $\mathbf{D}$ 是一个非零的信号，使 $\mathbf{d}=\mathcal{O}\mathbf{D}$ 的最大元素当 $N\to\infty$ 时是有界的，因此渐进地我们可以估计 $\mathbf{D}$ 为一个零向量(这是因为最终的 $\delta^{(u)}$变得大得使$|d_l+e_l|$不能超过它). 在判断统一阈值的用法的发展渐进证明过程中，如果我们成功地提取 $\mathbf{D}$，所有的非零信号变换系数当 $N\to\infty$ 时保持在 $\delta^{(u)}$的前面. 这需要"玩渐进游戏"，这在时间序列看来是很特别的方式. 作为例子(细节是练习[10.2]的主题)，我们假设有个时间序列 $\mathbf{X}$ 长度 $N=16$，包含信号 $\mathbf{D}$，它的值与哈尔离散小波变换矩阵的行 $n=14$ 成比例(参见图 57)：

$$D_t\equiv\begin{cases}-1, & 0\leqslant t\leqslant 7;\\ 1, & 8\leqslant t\leqslant 15.\end{cases}\tag{401a}$$

这个信号的哈尔离散小波变换由 $d_{14}=4=\sqrt{N}$，$d_l=0$，$l\neq 14$ 组成，因此这个信号被孤立到一个信号变换系数. 假设我们增加 N 到 2 的指数倍而趋于无穷大，即令 $N=2^J$，$J=5$，6，…. 如果我们决定周期地扩展原始信号，那么扩展的哈尔离散小波变换有 2^{J-1}个非零系数，每一个都等于 4(这样非零系数的比例因此就固定在 1/16). 这个扩展模仿我们观察多个信号通过时间通道的情况，这在信号处理中是很普遍的(即两个观察值之间的取样间隔 Δt 保持固定的，或者是单位的，并且我们观察的全部时间范围是当 N 增加的时候 $N\Delta t$ 扩展到整个时间). 我们能够证明 $\mathbf{P}[|O_l|\leqslant\delta^{(u)}]\to 1$ 当 $N\to\infty$对于所有的 l 成立，因此硬阈值信号估计 $\hat{\mathbf{D}}^{(ht)}$为零向量的概率很高，因此不能重构 $\mathbf{D}$. 另一方面，假设我们决定用"填充法"扩展信号：

$$D_t\equiv\begin{cases}-1, & 0\leqslant t\leqslant \frac{N}{2}-1;\\ 1, & \frac{N}{2}\leqslant t\leqslant N-1.\end{cases}\tag{401b}$$

这个扩展模拟将使我们在确定时间范围中获得一个信号的更多的精细取样值(即当 N 增加的时候，我们减小 Δt 使得 $N\Delta t$ 保持固定，例如说，等于 1)．这种策略在信号处理中并不通用，这里 Δt 通常设置使得到的奈奎斯特密度 $f_N = 1/(2\Delta t)$大于在信号中的频率(参见 10.6 节的评论与扩展中的第[2]条，取样方案更精细是很自然的)．这个信号的哈尔离散小波变换有一个非零系数，$d_{N-2} = \sqrt{N}$．我们现在证明当 $N\to\infty$ 时 $\boldsymbol{P}[|O_l|\leqslant\delta^{(u)}]\to 1$ 对所有的 $l\neq N-2$ 成立，而 $\boldsymbol{P}[|O_{N-2}|\leqslant\delta^{(u)}]\to 0$，因此$\hat{\boldsymbol{D}}^{(ht)}$任意接近于 $\boldsymbol{D}$ 的概率很高．关于应用$\hat{\boldsymbol{D}}^{(ht)}$有效性实践者只有 16 个原始点的时间序列可以从这些渐进证明，发现两个非常冲突的想法中看出一些．

虽然关于 $\delta^{(u)}$ 的解释有点让人担心，基于这个阈值的信号提取在某些确定的实际应用中有出众的应用效果，特别是由于它的内在简单性．我们将在 10.5 节和 10.7 节中给出一些例子．

• 交叉赋值(cv)

对于独立同分布高斯噪声情况下选择 δ 的另一个途径是著名的交叉赋值．这个方法在统计文献中广泛应用，特别是在非参数回归中选择光滑参数的一个自动方法(Green and Silverman，1994)．为适应当前的情况，我们可以用两种方式阐述交叉赋值(Nason，1996)．第一种方法，称为“两倍”交叉赋值，分裂 X_0，…，X_{N-1}组成奇数和偶数下标的下采样．给出一个阈值，我们分别用奇数下标(偶数下标)下采样来估计一个信号，然后插值和比较——用平方差——分别用偶数下标(奇数下标)下采样．我们选择阈值——称为$\hat{\delta}^{(tfcv)}$——作为一个最小平方差的和．第二个方法称为“漏一”交叉赋值，从 X_0，…，X_{N-1}开始，除去 X_t得到两个子序列，即 X_t前面的序列和后面的序列．给定一个阈值 δ，我们用这些子序列来产生对 X_t 的估计 $X_t^{(\delta)}$．对每一个 t 重复上面的过程．然后选择阈值——称为$\hat{\delta}^{(loocv)}$——$\delta>0$ 的值极小化

$$\sum_{t=0}^{N-1}(X_t^{(\delta)}-X_t)^2. \tag{402}$$

然而，在这个过程中与离散小波变换一样有两个问题．首先，J_0层局部离散小波变换只是自然地定义样本尺寸为 2^{J_0} 的倍数，然而现在这里我们有全长为 $N-1$ 的两个子序列．第二，我们需要某些特别的规则来定义 $X_t^{(\delta)}$．如我们在 10.5 节末尾中注意的，这些困难是可以克服的．

• 最小无偏风险

到目前为止，我们假设噪声变换系数为独立同分布高斯向量 $\boldsymbol{e}$，由变换一个相似的向量 ε 建模 $\boldsymbol{X}$ 作为“信号加噪声”的一部分组成．当 $\boldsymbol{X}$ 中的附加噪声替代为一个非独立同分布的向量 η 时，我们必须考虑噪声变换系数也是非独立同分布的．现在我们考虑对一个简单的非独立同分布模型的阈值选择大小，我们就直接

借助于噪声变换系数公式化．相应地，我们假设变换模型是

$$\mathbf{O} \equiv \mathcal{O}\mathbf{X} = \mathcal{O}\mathbf{D} + \mathcal{O}\eta \equiv \mathbf{d} + \mathbf{n},$$

所以

$$O_l = d_l + n_l,$$

其中 d_l 和 n_l 分别是 $\mathbf{d} \equiv \mathcal{O}\mathbf{D}$ 和 $\mathbf{n} \equiv \mathcal{O}\eta$ 的第 l 个分量．我们进一步假设

$$n_l \stackrel{\mathrm{d}}{=} \mathcal{N}(0, \sigma_{n_l}^2).$$

因此，每个分量 n_l 都是均值为零、方差为 $\sigma_{n_l}^2$ 的高斯分布，但是对噪声变换系数，模型不再是独立同分布，因为随机变量 n_l 允许有不同的方差，而且可以相关．

现在，令 $O_l^{(\delta)}$ 是 d_l 的一个估计，用来确定阈值 δ．令 $\mathbf{O}^{(\delta)}$ 是第 l 个分量为 $O_l^{(\delta)}$ 的向量．定义 $\hat{\mathbf{D}}^{(\delta)} \equiv \mathcal{O}^{\mathrm{T}}\mathbf{O}^{(\delta)}$ 并且注意到 $\mathcal{O}\hat{\mathbf{D}}^{(\delta)} = \mathcal{O}\mathcal{O}^{\mathrm{T}}\mathbf{O}^{(\delta)} = \mathbf{O}^{(\delta)}$．定义用 $\hat{\mathbf{D}}^{(\delta)}$ 估计 $\mathbf{D}$ 的风险 $R(\cdot, \cdot)$ 为

$$\begin{aligned} R(\hat{\mathbf{D}}^{(\delta)}, \mathbf{D}) \equiv E\{\|\hat{\mathbf{D}}^{(\delta)} - \mathbf{D}\|^2\} = E\{\|\mathcal{O}(\hat{\mathbf{D}}^{(\delta)} - \mathbf{D})\|^2\} &= E\{\|\mathbf{O}^{(\delta)} - \mathbf{d}\|^2\} \\ &= E\Big\{\sum_{l=0}^{N-1}(O_l^{(\delta)} - d_l)^2\Big\}, \end{aligned}$$

其中规范正交变换 $\mathcal{O}$ 不影响平方范数．按照 Stein(1981)，让我们考虑一类严格的估计，即可以写为

$$O_l^{(\delta)} = O_l + A^{(\delta)}(O_l), \quad l = 0, \cdots, N-1, \tag{403}$$

其中形式上 $A^{(\delta)}(\cdot)$ 必定是一个"弱可微"的实值函数(对这个概念的精确定义，参见 Härdle 等，1998，p. 72；严格地讲，我们希望 $A^{(\delta)}(\cdot)$ 至少是分段连续的)．那么我们有

$$O_l^{(\delta)} - d_l = n_l + A^{(\delta)}(O_l), \quad l = 0, \cdots, N-1,$$

因此

$$E\{(O_l^{(\delta)} - d_l)^2\} = \sigma_{n_l}^2 + 2E\{n_l A^{(\delta)}(O_l)\} + E\{[A^{(\delta)}(O_l)]^2\}.$$

既然 n_l 是均值为零、方差为 $\sigma_{n_l}^2$ 的高斯分布，我们有(进行变量替换 $\lambda \equiv d_l + n_l$)

$$\begin{aligned} E\{n_l A^{(\delta)}(O_l)\} &= E\{n_l A^{(\delta)}(d_l + n_l)\} \\ &= \frac{1}{\sqrt{(2\pi\sigma_{n_l}^2)}}\int_{-\infty}^{\infty} n_l A^{(\delta)}(d_l + n_l)\mathrm{e}^{-n_l^2/(2\sigma_{n_l}^2)}\mathrm{d}n_l \\ &= \frac{1}{\sqrt{(2\pi\sigma_{n_l}^2)}}\int_{-\infty}^{\infty}(\lambda - d_l)A^{(\delta)}(\lambda)\mathrm{e}^{-(\lambda-d_l)^2/(2\sigma_{n_l}^2)}\mathrm{d}\lambda \\ &= -\frac{1}{\sqrt{(2\pi\sigma_{n_l}^2)}}\int_{-\infty}^{\infty}\sigma_{n_l}^2 A^{(\delta)}(\lambda)\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathrm{e}^{-(\lambda-d_l)^2/(2\sigma_{n_l}^2)}\mathrm{d}\lambda. \end{aligned}$$

用分部积分

$$E\{n_l A^{(\delta)}(O_l)\} = \frac{1}{\sqrt{(2\pi\sigma_{n_l}^2)}}\int_{-\infty}^{\infty}\left[\sigma_{n_l}^2\frac{\mathrm{d}}{\mathrm{d}\lambda}A^{(\delta)}(\lambda)\right]\mathrm{e}^{-(\lambda-d_l)^2/(2\sigma_{n_l}^2)}\mathrm{d}\lambda,$$

这由弱可微性质得到. 因此

$$E\{n_l A^{(\delta)}(O_l)\} = \sigma_{n_l}^2 E\left\{\frac{\mathrm{d}}{\mathrm{d}\lambda}A^{(\delta)}(\lambda)\bigg|_{\lambda=O_l}\right\}.$$

现在，我们可以写

$$E\{(O_l^{(\delta)} - d_l)^2\} = E\{\mathcal{R}(\sigma_{n_l}, O_l, \delta)\}, \tag{404a}$$

其中$\mathcal{R}(\cdot, \cdot, \cdot)$的泛函形式是

$$\mathcal{R}(\sigma_{n_l}, x, \delta) \equiv \sigma_{n_l}^2 + 2\sigma_{n_l}^2 \frac{\mathrm{d}}{\mathrm{d}x}A^{(\delta)}(x) + [A^{(\delta)}(x)]^2.$$

使用等式(404a)，我们有

$$R(\hat{\mathbf{D}}^{(\delta)}, \mathbf{D}) = E\left\{\sum_{l=0}^{N-1}(O_l^{(\delta)} - d_l)^2\right\} = E\left\{\sum_{l=0}^{N-1}\mathcal{R}(\sigma_{n_l}, O_l, \delta)\right\},$$

其中右边的大括号里面的量我们称为对于$R(\hat{\mathbf{D}}^{(\delta)}, \mathbf{D})$的 Stein 的无偏风险估计(Stein，1981). 实际上，给出随机变量O_l的实现值o_l，$l=0, \cdots, N-1$，我们寻找使得下面式子最小的δ值：

$$\sum_{l=0}^{N-1}\mathcal{R}(\sigma_{n_l}, o_l, \delta).$$

随着δ基于 Stein 无偏风险估计而选择之后，我们可通过式(403)形成信号变换系数d_l的估计$O_l^{(\delta)}$.

作为一个适合式(403)结构的估计的一个例子，让我们考虑在式(400a)中定义的软阈值函数. 如果我们取$A^{(\delta)}(\cdot)$为分段连续函数

$$A^{(\delta)}(O_l) = \begin{cases} -O_l, & |O_l| < \delta; \\ -\delta\,\mathrm{sign}\{O_l\}, & |O_l| \geqslant \delta, \end{cases}$$

然后得到

$$O_l^{(\delta)} = O_l + A^{(\delta)}(O_l) = O_l^{(\mathrm{st})},$$

我们可以将这个$O_l^{(\delta)}$再写为

$$O_l^{(\delta)} = O_l - O_l 1_{[0,\delta^2)}(O_l^2) - \delta\,\mathrm{sign}\{O_l\} 1_{[\delta^2,\infty)}(O_l^2),$$

其中$1_{\mathcal{J}}(x)$表示指示函数

$$1_{\mathcal{J}}(x) \equiv \begin{cases} 1, & x \in \mathcal{J}; \\ 0, & \text{其他}. \end{cases}$$

练习[404] 证明，对于$A^{(\delta)}(O_l)$的这种选择

$$\mathcal{R}(\sigma_{n_l}, O_l, \delta) = O_l^2 - \sigma_{n_l}^2 + (2\sigma_{n_l}^2 - O_l^2 + \delta^2) 1_{[\delta^2,\infty)}(O_l^2). \tag{404b}$$

◁

现在，在式(404b)中的 $O_l^2-\sigma_{n_l}^2$ 不依赖于 δ，所以给定 $O_l=o_l$，$l=0$，…，$N-1$，δ 的 Stein 的无偏风险估计，即 $\delta^{(s)}$，是极小化下式给出的 δ 值：

$$\sum_{l=0}^{N-1}(2\sigma_{n_l}^2-o_l^2+\delta^2)1_{[\delta^2,\infty)}(o_l^2). \tag{405a}$$

假设我们对 $\{o_l\}$ 排序来给出 $\{o_{(l)}\}$ 使得

$$o_{(0)}^2\leqslant o_{(1)}^2\leqslant\cdots\leqslant o_{(N-1)}^2,$$

然后我们根据相应的 $\{o_l\}$ 的顺序来给出 $\{\sigma_{n_l}^2\}$ 的排序 $\{\sigma_{n_{(l)}}^2\}$. 为了找到 $\delta^{(s)}$，我们需要找到 δ^2 的平方根，使得下面的函数极小：

$$\Upsilon(\delta^2)\equiv\sum_{l=0}^{N-1}(2\sigma_{n_{(l)}}^2-o_{(l)}^2+\delta^2)1_{[\delta^2,\infty)}(o_{(l)}^2).$$

定义 $o_{(-1)}=0$ 使得 $o_{(k-1)}^2\leqslant o_{(k)}^2$ 对 $k=0$，…，$N-1$ 成立. 如果 δ 对某些 k 满足 $o_{(k-1)}^2<\delta^2<o_{(k)}^2$. 然后我们有

$$\begin{aligned}\Upsilon(\delta^2)&=(N-k)\delta^2+\sum_{l=k}^{N-1}(2\sigma_{n_{(l)}}^2-o_{(l)}^2)\\&>(N-k)o_{(k-1)}^2+\sum_{l=k}^{N-1}(2\sigma_{n_{(l)}}^2-o_{(l)}^2)=\Upsilon(o_{(k-1)}^2),\end{aligned}$$

因此 $\Upsilon(\cdot)$ 在 $o_{(k-1)}^2\leqslant\delta^2<o_{(k)}^2$ 的最小值一定在 $\delta^2=o_{(k-1)}^2$. 另一方面，如果 $\delta^2\geqslant o_{(N-1)}^2$，那么总有 $\Upsilon(\delta^2)=0$，因此我们选择 $\delta^2=o_{(N-1)}^2$ 在 δ^2 邻域内达到 $\Upsilon(\cdot)$ 的最小值. 因此我们能够通过计算 $\Upsilon(o_{(k)}^2)$，$k=-1$，0，…，$N-1$ 找到 $\Upsilon(\cdot)$ 在 δ^2 的最小值. 如果我们比较

$$\Upsilon(o_{(k)}^2)=(N-k-1)o_{(k)}^2+\sum_{l=k+1}^{N-1}(2\sigma_{n_{(l)}}^2-o_{(l)}^2)$$

与

$$\Upsilon(o_{(k+1)}^2)=(N-k-2)o_{(k+1)}^2+\sum_{l=k+2}^{N-1}(2\sigma_{n_{(l)}}^2-o_{(l)}^2),$$

就可以看到一个方便的迭代以 $\Upsilon(o_{(N-1)}^2)=0$ 开始，然后对于 $k=N-2$，$N-3$，…，0，-1 计算

$$\begin{aligned}\Upsilon(o_{(k)}^2)&=\Upsilon(o_{(k+1)}^2)+(2\sigma_{n_{(k+1)}}^2-o_{(k+1)}^2)+(N-k-1)o_{(k)}^2-(N-k-2)o_{(k+1)}^2\\&=\Upsilon(o_{(k+1)}^2)+2\sigma_{n_{(k+1)}}^2+(N-k-1)(o_{(k)}^2-o_{(k+1)}^2).\end{aligned} \tag{405b}$$

在一个简单的阈值水平 δ 中，刚刚描述的方法应用到所有分量 O_l，即使它们相关的噪声项 n_l 有不同的方差. 实际上，如果每个 n_l 对于所有的 l 有相同的方差 $\sigma_{n_l}^2=\sigma_0^2$，这个简单的基于 Stein 的无偏风险估计的阈值水平应该是 δ 的使得下式最

小的非负值：

$$\sum_{l=0}^{N-1}(2\sigma_0^2-o_l^2+\delta^2)1_{[\delta^2,\infty)}(o_l^2). \tag{406a}$$

• **具有共同方差的系数群**

对某些信号估计问题(例如参见 10.7 节)，噪声变换系数某个群具有一个共同的方差. 假设我们对分量 $\{O_l\}$ 重新换角标，使 $O_{j,t}$ 表示第 j 群的第 t 个系数，第 j 群记为 C_j，即

$$O_{j,t}=d_{j,t}+n_{j,t},\quad 其中\ n_{j,t}\stackrel{d}{=}\mathcal{N}(0,\sigma_j^2);$$

即第 j 群的成员有相同的方差 σ_j^2. 那么，对每个群 j 并给定 $O_{j,t}=o_{j,t}$. 从式(406a)得到一个阈值 $\delta_j^{(s)}$ 可以用下式最小的 δ 值来求得：

$$\sum_{t\in C_j}(2\sigma_j^2-o_{j,t}^2+\delta^2)1_{[\delta^2,\infty)}(o_{j,t}^2), \tag{406b}$$

其中，对于每个 j，我们可以用跟从前一样的算法. 定义一个随水平而变、基于 Stein 无偏风险估计的阈值，对软阈值有效.

统一阈值方法可以进行修改以适应这样的情况. 在信号 $\boldsymbol{D}$ 为零向量的假设下，变换系数跟噪声是相同的，即 $O_{j,t}=n_{j,t}$. 随机变量 $n_{j,t}$ 是相关的、高斯的，并且

$$\boldsymbol{P}[\max_{j,t}\{|n_{j,t}|/\sigma_j\}>\sqrt{(2\log(N))}]\leqslant B_c\leqslant\frac{1}{\sqrt{4\pi\log(N)}}$$

(Johnstone and Silverman，1997)，其中 B_c 表示用于相关的情况的一个界. 这个界是要比非独立情况下的界小. 进而，我们可以写

$$\boldsymbol{P}[\max_j\{\max_t\{|n_{j,t}|\}>\sqrt{(2\sigma_j^2\log(N))}\}]\leqslant B_c,$$

这样

$$\boldsymbol{P}[\max_j\{\max_t\{|n_{j,t}|\}\leqslant\delta_j^{(u)}]\to 1,$$

其中

$$\delta_j^{(u)}\equiv\sqrt{2\sigma_j^2\log(N)} \tag{406c}$$

对于 j 群的系数提供了一个保守的阈值水平，因此定义了一个随水平而变的“统一”阈值. 随着 $O_{j,t}$ 由 $o_{j,t}$ 的实现，$\delta_j^{(u)}$ 就是这样的集合. 现在，我们可以执行，例如，硬阈值(参见式(399))在每一个群 j 可以取

$$O_{j,t}^{(\mathrm{ht})}=\begin{cases}0, & |o_{j,t}|\leqslant\delta_j^{(u)};\\ o_{j,t} & 其他,\end{cases}\quad t\in C_j.$$

10.3 通过尺度方法的随机信号估计

假设 $\mathbf{X}=\mathbf{C}+\eta$，其中 $\mathbf{C}$ 表示一个感兴趣的未知的高斯信号，而 η 是不期望的零均值的高斯噪声，即 $\mathbf{C}$ 和 η 都是多元高斯随机变量，$E\{\eta\}=\mathbf{0}$，并且协方差矩阵分别是 Σ_C 和 Σ_η. 我们假设随机信号 $\mathbf{C}$ 与 η 是不相关的(即 C 中任意随机变量和 η 中的随机变量之间的协方差为零——回想两个非相关的高斯随机变量实际上是独立的).

现在，令 $\mathcal{O}$ 为一个 $N\times N$ 阶规范正交矩阵，并且令 $\mathbf{O}\equiv\mathcal{O}\mathbf{X}$ 表示相对应于 $\mathbf{X}$ 的变换系数 $\{O_l: l=0, \cdots, N-1\}$ 的 N 维向量. 那么

$$\mathbf{O}\equiv\mathcal{O}\mathbf{X}=\mathcal{O}\mathbf{C}+\mathcal{O}\eta\equiv\mathbf{R}+\mathbf{n}, \text{因此 } O_l=R_l+n_l,$$

其中 R_l 和 n_l 分别是 $\mathbf{R}\equiv\mathcal{O}\mathbf{C}$ 和 $\mathbf{n}\equiv\mathcal{O}\eta$ 的第 l 个分量. 由多元高斯随机变量理论的一个标准结果表明，$\mathbf{R}$ 和 $\mathbf{n}$ 分别具有协方差矩阵 $\mathcal{O}\Sigma_C\mathcal{O}^T$ 和 $\mathcal{O}\Sigma_\eta\mathcal{O}^T$ 的高斯随机变量(参见式(262c))；进而，因为 $\mathbf{C}$ 和 η 是不相关的，他们的变换 $\mathbf{R}$ 和 $\mathbf{n}$ 也是如此，这暗示 R_l 和 $n_{l'}$ 对所有的 l 和 l' 是不相关的.

因为 η 有零均值，从式(262b)我们知道 $E\{\mathbf{n}\}=0$，特别地 n_l 具有零均值. 这里我们假设任意一个感兴趣的分量 l 使得 $E\{R_l\}=0$. 注意我们不需要 $E\{\mathbf{C}\}=0$，因为变换 $\mathcal{O}$ 可以对于许多，而不是全部的系数能够达到. (回想，例如，离散小波变换限制尺度系数中时间序列的取样平均，这些尺度系数作为全部离散小波变换系数的一小部分). 因此，在本章中，当观察统计信号估计时，我们假设处理的分量有 $E\{R_l\}=0$. 如 10.5 节将指出的，这对于离散小波变换是容易获得的.

假设我们希望估计信号系数 R_l 而借助于观察的信号加噪声系数 O_l，使用一个简单的加权值 a_l，即

$$\hat{R}_l\equiv a_l O_l. \tag{407}$$

附加假设，我们决定选择 a_l 使得 $E\{(R_l-\hat{R}_l)^2\}$ 尽可能地小.

练习[407] 证明极小化 $E\{(R_l-a_lO_l)^2\}$ 要求 $E\{(R_l-a_lO_l)O_l\}=0$，这是著名的线性最小二乘估计的正交性法则. ◁

总之，估计误差和估计量彼此不相关(即“正交的”). 因此

$$E\{R_lO_l\}-a_lE\{O_l^2\}=0, \quad \text{所以 } a_l=\frac{E\{R_lO_l\}}{E\{O_l^2\}}.$$

由 $\{R_l\}$ 和 $\{n_l\}$ 是不相关的和高斯的事实，我们有

$$E\{R_lO_l\}=E\{R_l^2\} \quad \text{和} \quad E\{O_l^2\}=E\{R_l^2\}+E\{n_l^2\}.$$

因此

$$\hat{R}_l=\frac{E\{R_l^2\}}{E\{R_l^2\}+E\{n_l^2\}}O_l.$$

因为随机变量 R_l 和 n_l 有零均值，数量 $\sigma_{R_l}^2 \equiv E\{R_l^2\}$ 和 $\sigma_{n_l}^2 \equiv E\{n_l^2\}$ 分别是信号和噪声的方差. 因此，信号分量 R_l 的估计 $\hat{R}_l$ 可写作观察系数 O_l 用信号方差与全方差的比加权的第 l 个分量：

$$\hat{R}_l = \frac{\sigma_{R_l}^2}{\sigma_{R_l}^2 + \sigma_{n_l}^2} O_l. \tag{408}$$

在第 l 个分量中噪声出现的时候，这个权值有效地将 O_l"收缩"到零，而且实际上如果噪声水平很高，可能完全"扼杀"O_l，传递一个接近于零的估计 $\hat{R}_l$. 在第 l 个分量中没有噪声的情况下，系数 O_l 返回作为 R_l 的估计，直觉上是有道理的.

10.4 通过收缩方法的随机信号估计

在 10.2 节中我们看到了阈值法则，即设某些变换系数恰好是零，而处理其他的系数为不变的或者是减小幅值. 在本节中，我们集中于收缩法则. 其中不同于阈值法则在于所有的非零的变换系数被映射为非零值——实际上，它们或者在某些光滑的意义下非线性减少幅值或者保持不变. 然而，实际上收缩法则常常像阈值法则，因为有个确定范围的非零系数收缩到非常接近零.

• 条件均值和中值

我们继续 10.3 节中的思想，把信号和噪声作为独立的统计分量，但是注意到我们没有先验假设：信号和噪声是高斯随机变量. 规范正交变换采用形式

$$\boldsymbol{O} \equiv \mathcal{O}\boldsymbol{X} = \mathcal{O}\boldsymbol{C} + \mathcal{O}\eta \equiv \boldsymbol{R} + \boldsymbol{n}, \quad \text{因此 } O_l = R_l + n_l,$$

其中 R_l 是被变换的统计信号 $\boldsymbol{R}$ 的第 l 个分量，n_l 是被变换的噪声 $\boldsymbol{n}$ 的第 l 个分量，并且我们假设 $E\{R_l\} = E\{n_l\} = 0$. 在 10.3 节中我们看到，如果我们希望产生基于尺度 O_l 的 R_l 的一个线性估计，而且如果我们用均方误差标准，那么尺度因之是信号方差和整体方差的比来作为其第 l 个分量. 假设我们仍然用均方误差标准，但我们现在考虑一个更为广泛的(非线性的)估计，R_l 的估计 $U_2(O_l)$ 就是 $\hat{R}_l = U_2(O_l)$，而 $U_2(O_l)$ 定义了运算于 O_l 的收缩法则. 像我们将证明的是使得 $E\{(R_l - U_2(O_l))^2\}$ 极小的 R_l 的估计 $U_2(O_l)$ 是条件均值 $E\{R_l \mid O_l\}$(下标"2"强调这是一个平方范数的极小值).

相应地，令 $\boldsymbol{U}_2$ 是第 l 个分量为 $U_2(O_l)$ 的向量. 因为 $U_2(O_l)$ 是 R_l 的一个估计，向量 $\boldsymbol{U}_2$ 是 $\boldsymbol{R} = \mathcal{O}\boldsymbol{C}$ 的估计，即 $\boldsymbol{R}$ 是变换后的统计信号. 定义 $\hat{\boldsymbol{C}} \equiv \mathcal{O}^{\mathrm{T}}\boldsymbol{U}_2$，并且注意 $\mathcal{O}\hat{\boldsymbol{C}} = \boldsymbol{U}_2$. 涉及用 $\hat{\boldsymbol{C}}$ 估计 $\boldsymbol{C}$ 的风险 $R(\cdot, \cdot)$ 由下式给出：

$$\begin{aligned} R(\hat{\boldsymbol{C}}, \boldsymbol{C}) &\equiv E\{\|\hat{\boldsymbol{C}} - \boldsymbol{C}\|^2\} = E\{\| \mathcal{O}(\hat{\boldsymbol{C}} - \boldsymbol{C})\|^2\} \\ &= E\{\|\boldsymbol{U}_2 - \boldsymbol{R}\|^2\} \\ &= E\left\{\sum_{l=0}^{N-1}(U_2(O_l) - R_l)^2\right\} = \sum_{l=0}^{N-1} E\{(U_2(O_l) - R_l)^2\}, \end{aligned}$$

其中我们再次使用事实规范正交变换$\mathcal{O}$不影响平方范数. 因此, 这个风险极小化明显通过 $E\{(U_2(O_l)-R_l)^2\}$ 对每个 l 极小化得到. 我们声明, R_l 的估计 $U_2(O_l)$使得这个平方误差标准极小的是条件均值 $E\{R_l \mid O_l\}$. 为此注意

$$E\{(U_2(O_l)-R_l)^2\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(U_2(o_l)-r_l)^2 f_{R_l,O_l}(r_l,o_l)\mathrm{d}r_l\mathrm{d}o_l,$$

其中 $f_{R_l,O_l}(\cdot,\cdot)$是随机变换 R_l 和O_l的联合概率密度函数(PDF). 事件的条件是 O_l假设取值o_l, 我们可以写成 $f_{R_l,O_l}(r_l, o_l)=f_{R_l\mid O_l=o_l}(r_l)f_{O_l}(o_l)$, 这就产生

$$E\{(U_2(O_l)-R_l)^2\}=\int_{-\infty}^{\infty}f_{O_l}(o_l)\int_{-\infty}^{\infty}(U_2(o_l)-r_l)^2 f_{R_l\mid O_l=o_l}(r_l)\mathrm{d}r_l\mathrm{d}o_l.$$

里面的被积函数是非负的, $f_{O_l}(o_l)$也是, 所以二重积分的极小值可通过对每一个结果 o_l使

$$\int_{-\infty}^{\infty}(U_2(o_l)-r_l)^2 f_{R_l\mid O_l=o_l}(r_l)\mathrm{d}r_l$$

极小获得. 如果我们在式(261a)中分别用 O_l 和 R_l替换 X_0 和 X_1, 那么我们看到, 上面的积分等同于式(261a)中的积分, 因此

$$E\{R_l \mid O_l=o_l\}=\arg\min_{U_2(o_l)}\int(U_2(o_l)-r_l)^2 f_{R_l\mid O_l=o_l}(r_l)\mathrm{d}r_l,$$

如要求的那样.

现在, 条件概率密度函数 $f_{R_l\mid O_l=o_l}(r_l)$可以写为

$$f_{R_l\mid O_l=o_l}(r_l)=\frac{f_{R_l,O_l}(r_l,o_l)}{f_{O_l}(o_l)}.$$

从练习[262a], 我们知道

$$f_{R_l,O_l}(r_l,o_l)=f_{R_l,n_l}(r_l,o_l-r_l),$$

其中 $f_{R_l,n_l}(\cdot,\cdot)$是随机变量 R_l 和n_l的联合概率密度函数. 因为我们已经假设 $\mathbf{C}$ 和 η 是独立的, 它们的变换 $\mathbf{R}=\mathcal{O}\,\mathbf{C}$ 和 $\mathbf{n}=\mathcal{O}\,\eta$ 也必定是独立的. 由于 R_l 和 n_l 因此是独立的, 我们有 $f_{R_l,n_l}(r_l, o_l-r_l)=f_{R_l}(r_l)f_{n_l}(o_l-r_l)$. 因此, 对独立于信号的附加噪声, 我们得到

$$f_{R_l,O_l}(r_l,o_l)=f_{R_l}(r_l)f_{n_l}(o_l-r_l).$$

进而, 因为 $O_l=R_l+n_l$, O_l 的概率密度函数是一个和, 并且我们可使用这个结果(参见式(262a)): 独立随机变量的和是它们密度的卷积, 得到

$$f_{O_l}(o_l)=f_{R_l}*f_{n_l}(o_l)=\int_{-\infty}^{\infty}f_{R_l}(r_l)f_{n_l}(o_l-r_l)\mathrm{d}r_l,$$

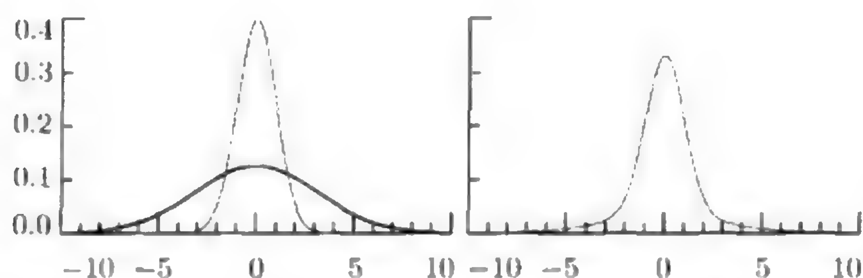


图 410 随机变量$\mathcal{N}(0, 1)$和$\mathcal{N}(0, 10)$的概率密度函数(左边图形，分别是细和粗曲线)，一个遵守混合高斯模型的随机变量(右边图形)．混合概率密度函数是非高斯的，是由$\mathcal{N}(0,1)$和$\mathcal{N}(0, 10)$的概率密度函数的加权形成的，权值分别为 $p_l=0.75$ 和 $1-p_l=0.25$(取自 Chipman 等，1997 的图 1)

因此

$$f_{R_l|O_l=o_l}(r_l)=\frac{f_{R_l}(r_l)f_{n_l}(o_l-r_l)}{\int_{-\infty}^{\infty}f_{R_l}(r_l)f_{n_l}(o_l-r_l)\mathrm{d}r_l}. \tag{410a}$$

给定 O_l 的取值 o_l，我们能够将条件均值写为

$$E\{R_l\mid O_l=o_l\}=\int_{-\infty}^{\infty}r_lf_{R_l|O_l=o_l}(r_l)\mathrm{d}r_l=\frac{\int_{-\infty}^{\infty}r_lf_{R_l}(r_l)f_{n_l}(o_l-r_l)\mathrm{d}r_l}{\int_{-\infty}^{\infty}f_{R_l}(r_l)f_{n_l}(o_l-r_l)\mathrm{d}r_l}. \tag{410b}$$

现在，让我们看一种特殊情况，统计信号遵守混合高斯模型，如下定义．令 $\mathcal{I}_l$ 是一个二值的随机变量，使

$$\mathbf{P}[\mathcal{I}_l=1]=p_l \quad 和 \quad P[\mathcal{I}_l=0]=1-p_l, \tag{410c}$$

其中 $0\leqslant p_l\leqslant 1$．我们说 R_l 遵循混合高斯模型，如果

$$R_l\stackrel{\mathrm{d}}{=}\mathcal{I}_l\,\mathcal{N}(0,\gamma_l^2\sigma_{G_l}^2)+(1-\mathcal{I}_l)\,\mathcal{N}(0,\sigma_{G_l}^2), \tag{410d}$$

其中，高斯随机变量$\mathcal{N}(0, \gamma_l^2\sigma_{G_l}^2)$和$\mathcal{N}(0, \sigma_{G_l}^2)$是相互独立的，$\mathcal{I}_l$ 也是独立的．我们可以通过首先产生$\mathcal{I}_l$的值，然后产生$\mathcal{N}(0, \gamma_l^2\sigma_{G_l}^2)$如果$\mathcal{I}_l=1$，或者$\mathcal{N}(0, \sigma_{G_l}^2)$如果$\mathcal{I}_l=0$ 以得到 R_l的值．这个统计信号是两个均值为零、方差分别为 $\gamma_l^2\sigma_{G_l}^2$ 和 $\sigma_{G_l}^2$、比例为 p_l比 $1-p_l$的高斯分布的混合．混合高斯模型是一种所有非高斯信号的建模方法．例如，如果 $\gamma_l\ll 1$ 并且 $p_l=0.95$，那么我们允许信号项的 5%方差 $\sigma_{G_l}^2$ 比剩下的 95%的方差 $\gamma_l^2\sigma_{G_l}^2$ 大．R_l 的概率密度函数是由 $\mathcal{N}(0, \gamma_l^2\sigma_{G_l}^2)$和$\mathcal{N}(0, \sigma_{G_l}^2)$的概率密度函数权值为 p_l 和 $1-p_l$的加权和．图 410 中显示了一个例子，其中 $p_l=0.75$，$\gamma_l^2\sigma_{G_l}^2=1$，$\sigma_{G_l}^2=10$.

对此特殊情形，为了完成对我们的模型，我们假设噪声是均值为零、方差 $\sigma_{n_l}^2$ 的高斯型噪声：

$$n_l\stackrel{\mathrm{d}}{=}\mathcal{N}(0,\sigma_{n_l}^2). \tag{410e}$$

练习[10.5]表明，给定这个准备，我们有

$$E\{R_l\mid O_l=o_l\}=\frac{a_lA_l(o_l)+b_lB_l(o_l)}{A_l(o_l)+B_l(o_l)}o_l, \tag{410f}$$

其中

$$a_l \equiv \frac{\gamma_l^2\sigma_{G_l}^2}{\gamma_l^2\sigma_{G_l}^2+\sigma_{n_l}^2};\quad A_l(o_l) \equiv \frac{p_l}{\sqrt{(2\pi[\gamma_l^2\sigma_{G_l}^2+\sigma_{n_l}^2])}}\mathrm{e}^{-o_l^2/[2(\gamma_l^2\sigma_{G_l}^2+\sigma_{n_l}^2)]};$$

$$b_l \equiv \frac{\sigma_{G_l}^2}{\sigma_{G_l}^2+\sigma_{n_l}^2};\quad 并且\ B_l(o_l) \equiv \frac{1-p_l}{\sqrt{(2\pi[\sigma_{G_l}^2+\sigma_{n_l}^2])}}\mathrm{e}^{-o_l^2/[2(\sigma_{G_l}^2+\sigma_{n_l}^2)]}. \tag{411a}$$

因为 $U_2(O_l)=E\{R_l \mid O_l\}$，$R_l$的条件均值估计 $U_2(O_l)$现在按照 O_l是非线性的，因为 O_l^2 出现在表达式 $A_l(\cdot)$和 $B_l(\cdot)$中的指数中，其中这些都被认为是随机变量．式(410f)中的条件均值的形式将在本节的评论与扩展中进一步讨论．

为了获取对给出的式(410f)怎样收缩 o_l的一个深入的看法，让我们考虑一种特别重要的情况 $\gamma_l=0$(我们可以将$\mathcal{N}(0,0)$解释为退化的随机变量，其实现总为零)．信号的模型变为

$$R_l \stackrel{\mathrm{d}}{=} (1-\mathcal{I}_l)\mathcal{N}(0,\sigma_{G_l}^2), \tag{411b}$$

因此 $\boldsymbol{P}[R_l=0]=\boldsymbol{P}[\mathcal{I}_l=1]=p_l$．对于 p_l 的值接近于 1，这就给出了所谓的稀疏信号模型，即在很多情况下信号是零，只有偶尔是非零的．这个稀疏信号模型导致了条件均值的一个特别简单的形式

$$E\{R_l \mid O_l=o_l\}=\frac{b_l}{1+c_l}o_l, \tag{411c}$$

其中

$$c_l=\frac{p_l\sqrt{(\sigma_{G_l}^2+\sigma_{n_l}^2)}}{(1-p_l)\sigma_{n_l}^2}\mathrm{e}^{-o_l^2b_l/(2\sigma_{n_l}^2)}. \tag{411d}$$

我们注意到 $E\{R_l \mid O_l=o_l\}\approx o_lb_l$ 对于大的 o_l 成立，即由式(407)的线性形式族．这种条件均值的形式已经很成功地应用在地球物理中(Godfred and Rocca，1981；Walden，1985)．实际上，对于稀疏信号模型，为了计算收缩法则 $U_2(o_l)=E\{R_l \mid O_l=o_l\}$，需要三个参数值，即 p_l，$\sigma_{G_l}^2$和$\sigma_{n_l}^2$．另一方面，如果我们想对这个模型用 Stein 无偏风险估计软阈值，则只需噪声方差 $\sigma_{n_l}^2$——本质上需要处理三个参数来决定式(411c)中的收缩法则，是由于我们没有预先选定阈值函数的图形的这个事实．

图 412 解释了式(411c)的条件均值收缩法则，其中 $p_l=0.95$，$\sigma_{n_l}^2=1$，$\sigma_{G_l}^2=5$，10 和 25(分别从粗线到细线)．注意对于 $\sigma_{G_l}^2=25$，收缩法则类似于严格阈值法则的光滑形式(参见式(400c))．

条件均值估计 $U_2(\cdot)$是从极小化均方误差得来的．函数 $U_2(\cdot)$不是一个阈值法则，而是一个(非线性)收缩法则．在这个过程中没有系数被完全扼杀．我们同样可以很好地选择极小化均值绝对误差．在这种情况下，估计 $U_1(\cdot)$是条件均值(参见式(261b))．当 O_l 取值 o_l 时，我们需要对 $U_1(o_l)$解如下方程：

$$\int_{-\infty}^{U_1(o_l)} f_{R_l|O_l=o_l}(r_l)\mathrm{d}r_l=\frac{\int_{-\infty}^{U_1(o_l)} f_{R_l}(r_l)f_{n_l}(o_l-r_l)\mathrm{d}r_l}{\int_{-\infty}^{\infty} f_{R_l}(r_l)f_{n_l}(o_l-r_l)\mathrm{d}r_l}=0.5, \tag{411e}$$

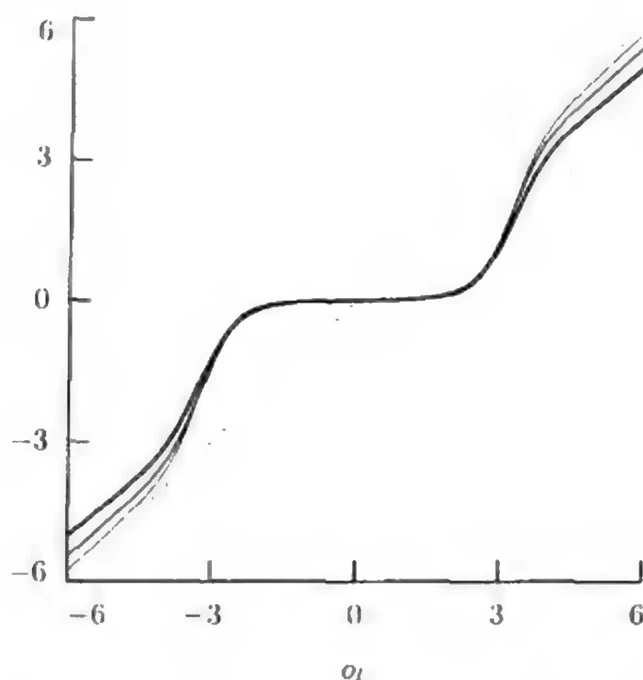


图412 式(411c)对 $p_l=0.95$，$\sigma_{n_l}^2=1$ 和 $\sigma_{G_l}^2=5$(最粗的曲线，最远离点斜线)的条件均值收缩法则，10 和 25(最细曲线，最接近斜线). 因为在条件均值收缩法则和 Chipman 等(1997)的关于平方误差损失贝叶斯估计法则之间的相对应性，上面的图也解了式(414a)的 $B_2(\cdot)$.

其中我们用了式(410a). Godfrey and Rocca(1981)指出，对于稀疏信号模型，收缩法则 $U_1(\cdot)$近似地等于一个硬阈值法则，即

$$U_1(O_l)\approx\begin{cases}0, & |O_l|\leqslant\delta;\\ b_lO_l, & \text{其他},\end{cases}\tag{412a}$$

其中阈值是由下式给出的：

$$\delta=\sigma_{n_l}\left[2\log\left(\frac{p_l\sigma_{G_l}}{(1-p_l)\sigma_{n_l}}\right)\right]^{1/2}\tag{412b}$$

(假设 $p_l/(1-p_l)\gg\sigma_{n_l}^2/(\sigma_{G_l}\delta)$ 和 $\sigma_{G_l}^2\gg\sigma_{n_l}^2$ 逼近是有效的).

• *贝叶斯方法*

很多最近的关于收缩的工作是建立在小波变换和贝叶斯方法基础上的(后者在 7.3 节中有简单的回顾). 一些因此而来的收缩法则可以从我们讨论的条件均值和中值中看出来. 我们考虑的这个讨论调整为假设信号是统计的，噪声不必是独立同分布的，因此我们写 $\mathbf{X}=\mathbf{C}+\boldsymbol{\eta}$. 应用规范正交变换 $\mathcal{O}$，我们获得模型 $\mathbf{O}=\mathbf{R}+\mathbf{n}$，我们将这个模型的第 l 个分量写为 $O_l=R_l+n_l$，对任一给定的 l，假设随机变量 R_l 和 n_l是相互独立的. 注意我们没有 $\mathbf{n}$ 中的随机变量是彼此相关的($\mathbf{R}$ 中的随机变量也是一样的)的任何假设. 我们不需要这样做是因为下面我们分别讨论每个信号分量的提取方案，即我们仅用 O_l估计信号分量 R_l. 然而，实际应用

这些算法需要估计确定的参数，作为一个整体来说正是这点我们需要具体指出 $\mathbf{n}$ 的模型结构(我们将在 10.5 节转到这点).

Chipman 等(1997)假设噪声变换系数 n_l，如式(410e)，是均值为零、方差为 $\sigma_{n_l}^2$ 的高斯噪声，并且信号变换系数 R_l 有一个混合高斯的概率密度函数，如式(410d)中指定的. 在贝叶斯公式中，我们把这个混合模型写为

$$R_l \mid \mathcal{I}_l \stackrel{d}{=} \mathcal{I}_l \mathcal{N}(0,\gamma_l^2\sigma_{G_l}^2)+(1-\mathcal{I}_l)\mathcal{N}(0,\sigma_{G_l}^2), \tag{413a}$$

其中$\mathcal{I}_l$ 被认为是一个“混合参数”，其先验分布由式(410c)决定. $\mathcal{I}_l$的分布依赖于 p_l，现在称为“超参数”，另外，随机变量 $R_l \mid \mathcal{I}_l$ 依赖于这个参数和其他两个超参数 γ_l^2 和$\sigma_{G_l}^2$. 注意 R_l 自己的概率密度函数(即没有条件的随机变量$\mathcal{I}_l$)是由 $\mathcal{N}(0, \gamma_l^2\sigma_{G_l}^2)$和$\mathcal{N}(0, \sigma_{G_l}^2)$概率密度函数加起来组成的，权值分别为 p_l 和 $1-p_l$. 如我们可在比较式(410d)和(413a)的右边看出来的，当我们没有接近实现$\mathcal{I}_l$的值的时候，在随机变量 R_l 和 $R_l \mid \mathcal{I}_l$之间没有明显的区别(当这些是可用的，那么给定$\mathcal{I}_l=1$，按分布 R_l等于随机变量$\mathcal{N}(0, \gamma_l^2\sigma_{G_l}^2)$而给定$\mathcal{I}_l=0$ 时，按分布 R_l等于$\mathcal{N}(0, \sigma_{G_l}^2)$). 因为 $O_l=R_l+n_l$，并且因为 n_l是高斯分布，均值为零，方差为 $\sigma_{n_l}^2$，即满足

$$O_l \mid R_l \stackrel{d}{=} \mathcal{N}(R_l,\sigma_{n_l}^2). \tag{413b}$$

实际上，为了获得现在称为“贝叶斯收缩法则”的方法，$\sigma_{n_l}^2$ 和三个超参数不得不估计.

Vidakovic(1998)也观察了贝叶斯收缩法则. 类似于 Chipman 等(1997)，他也假设 n_l是均值为零、方差为 $\sigma_{n_l}^2$ 的高斯分布，但是现在只有条件：他不认为方差 $\sigma_{n_l}^2$已知，而给出它的先验分布，即指数的超参数为 v_l. 因此代替式(410e)，现在我们有

$$n_l \mid \sigma_{n_l}^2 \stackrel{d}{=} \mathcal{N}(0,\sigma_{n_l}^2),$$

这转而指出

$$O_l \mid R_l,\sigma_{n_l}^2 \stackrel{d}{=} \mathcal{N}(R_l,\sigma_{n_l}^2). \tag{413c}$$

如 7.3 节所讨论的，高斯分布的混合方差是方差有一个指数分布乘拉普拉斯(两面指数的)分布，使对于 n_l 的边缘概率密度函数为

$$f_{n_l}(n_l)=\frac{1}{\sqrt{(2/v_l)}}\mathrm{e}^{-(2v_l)^{1/2}|n_l|}$$

(参见式(257c)). 同样的结论对 $O_l \mid R_l$ 也成立，所以我们也有

$$f_{O_l|R_l=r_l}(o_l)=\frac{1}{\sqrt{(2/v_l)}}\mathrm{e}^{-(2v_l)^{1/2}|o_l-r_l|}. \tag{413d}$$

代替式(413a)，Vidakovic(1998)完成了模型的详细说明，他给出 R_l 是一个 t 分布(式(258a))，自由度为 ϑ_l，尺度函数为 κ_l(位置参数被设置为零)．这个模型的超参数为 v_l，ϑ_l 和 κ_l．

令 $B_2(O_l)$ 为 R_l 关于平方误差损失的贝叶斯法则估计，即 $B_2(O_l)$ 是 $E_{R_l}\{E_{O_l R_l}\{(B_2(O_l)-R_l)^2\}\}$ 的极小化．如式(265b)所陈述的，贝叶斯法则在这种情况下是由后验均值给出的．给定 O_l 的实现 o_l，后验均值与式(410b)有相同的条件均值，即

$$B_2(o_l)=\int_{-\infty}^{\infty} r_l f_{R_l|O_l=o_l}(r_l)\,\mathrm{d}r_l,$$

并且 $B_2(o_l)\equiv U_2(o_l)$．现在

$$f_{R_l|O_l=o_l}(r_l)=\frac{f_{R_l,O_l}(r_l,o_l)}{f_{O_l}(o_l)}=\frac{f_{R_l}(r_l)f_{O_l|R_l=r_l}(o_l)}{f_{O_l}(o_l)}=\frac{f_{R_l}(r_l)f_{O_l|R_l=r_l}(o_l)}{\int_{-\infty}^{\infty} f_{R_l}(r_l)f_{O_l|R_l=r_l}(o_l)\mathrm{d}r_l},$$

因此我们能将 $B_2(o_l)$ 写为

$$B_2(o_l)=\frac{\int_{-\infty}^{\infty} r_l f_{R_l}(r_l)f_{O_l|R_l=r_l}(o_l)\mathrm{d}r_l}{\int_{-\infty}^{\infty} f_{R_l}(r_l)f_{O_l|R_l=r_l}(o_l)\mathrm{d}r_l}. \tag{414a}$$

对于 Chipman 等(1997)的模型，得到的后验均值(即贝叶斯法则)恒等于式(410f)给出的条件均值(因此在图 412 中说明)．接着是 $f_{O_l|R_l=r_l}(o_l)$ 由(413b)具体指出为

$$f_{O_l|R_l=r_l}(o_l)=\frac{1}{\sqrt{(2\pi\sigma_{n_l}^2)}}\mathrm{e}^{-(o_l-r_l)^2/(2\sigma_{n_l}^2)},$$

这恒等于式(410b)中的 $f_{n_l}(o_l-r_l)$(n_l 的概率密度函数在式(410e)中详细说明了)．R_l 的概率密度函数在两种情况下是相同的，并且在式(410d)中详细说明．

对于 Vidakovic(1998)的模型，我们通过信号的概率密度函数 $f_{R_l}(\cdot)$ 的 t 分布和对于 $f_{O_l|R_l=r_l}(o_l)$ 的拉普拉斯分布(如式(413d)得到的)得到式(414a)的贝叶斯法则．得到的 $B_2(\cdot)$ 不能降为可处理的形式，因此我们必须求助于数值积分．为了计算所需要的积分，我们注意到对于 $o_l\geqslant 0$，有

$$B_2(o_l)=\kappa_l\sqrt{\vartheta_l}\,\frac{A_1^-(o_l',o_l',\infty)+A_1^+(o_l',0,o_l')-\mathrm{e}^{-o_l\sqrt{(2v_l)}}A_1^-(0,0,\infty)}{A_0^-(o_l',o_l',\infty)+A_0^+(o_l',0,o_l')+\mathrm{e}^{-o_l\sqrt{(2v_l)}}A_0^-(0,0,\infty)}, \tag{414b}$$

其中 $o_l'\equiv o_l/(\kappa_l\sqrt{\vartheta_l})$，并且

$$A_n^{\pm}(o_l',a,b)\equiv\int_a^b \frac{x^n}{(1+x^2)^{(\vartheta_l+1)/2}}\mathrm{e}^{\pm(x-o_l')\kappa_l\sqrt{(2v_l\vartheta_l)}}\mathrm{d}x$$

(上面的证明是练习[10.6]；注意对于 $o_l<0$，我们能用关系式 $B_2(o_l)=-B_2(-o_l)$ 得到 $B_2(o_l)$．图 427 展示了小波系数而成的 $B_2(\cdot)$ 的两个例子．

如果我们令 $B_1(O_l)$ 是 R_l 关于绝对误差损失意义下的贝叶斯法则估计，即 $B_1(O_l)$ 是 $E_{R_l}\{E_{O_l|R_l}\{|B_1(O_l)-R_l|\}\}$ 的极小化．那么贝叶斯法则与后验中值（式(266)）是一样的．再次，给出 O_l 的实现 o_l，后验中值与式(411e)的条件中值是相同的，即是

$$\int_{-\infty}^{B_1(o_l)} f_{R_l|O_l=o_l}(r_l)\,\mathrm{d}r_l = 0.5$$

的解．因此 $B_1(o_l)\equiv U_1(o_l)$，并且在我们有附加噪声的、信号独立的假设下，一个可计算的形式在式(411e)的右边给出（式(412a)给出了一个方便的逼近）．Abramovich 等(1998)检查对于式(411b)的稀疏信号模型的后验中值．他们发现

$$B_1(o_l) = \mathrm{sign}\{o_l\}\max\{0,\zeta_l\}, \tag{415}$$

其中

$$\zeta_l = |o_l|b_l - (b_l\sigma_{n_l}^2)^{1/2}\Phi^{-1}\left(\frac{1+\min\{c_l,1\}}{2}\right),$$

并且 c_l 在式(411d)中定义（这里 $\Phi^{-1}(\cdot)$ 是标准高斯累积分布函数的逆——参见式(257b)）．对所有 o_l 在某一熟悉的定义区间 $[-\delta,\delta]$ 上，量 ζ_l 是负的．因此，当 $|o_l|$ 落在阈值 δ 的下面时，$B_1(o_l)$ 置为零，并且我们得到了一个阈值法则而不是一个收缩法则．

图 416 比较了由式(415)计算的后验中值（细曲线）和式(412a)（粗线）中的逼近条件均值，其中 $\sigma_{G_l}^2=25$，$p_l=0.95$，$\sigma_{n_l}^2=1$，如 Abramovich 等(1998)使用的．很清楚，实际上在这两种阈值法则之间的区别很少．因此，Godfrey and Rocca (1981)的逼近对于这种情况已经足够了．注意，这两个条件对于这个逼近是有效的，并且对于这里选择的参数是满足的．

10.4 节的评论与扩展

[1]这里我们发展一个对于式(410f)中的条件均值形式的解释．由式(410d)看来，我们可以把 R_l 表示为

$$R_l = \mathcal{I}_l R_l^{(1)} + (1-\mathcal{I}_l)R_l^{(2)},$$

其中的 $R_l^{(1)}$ 和 $R_l^{(2)}$ 是独立的高斯随机变量，均值为零而方差分别为 $\gamma_l^2\sigma_{G_l}^2$ 和 $\sigma_{G_l}^2$．令 $f_{R_l}(\cdot)$，$f_{R_l^{(1)}}(\cdot)$ 和 $f_{R_l^{(2)}}(\cdot)$ 分别是对于 R_l，$R_l^{(1)}$ 和 $R_l^{(2)}$ 的概率密度函数．因为 R_l 服从高斯混合模型，如前声明的，我们有

$$f_{R_l}(r_l) = p_l f_{R_l^{(1)}}(r_l) + (1-p_l)f_{R_l^{(2)}}(r_l).$$

回想时间序列的变换系数 O_l 可以写为 $O_l=R_l+n_l$，其中 n_l 是均值为零、方差为 $\sigma_{n_l}^2$ 的高斯随机变量，独立于 R_l．令 $f_{O_l}(\cdot)$ 和 $f_{n_l}(\cdot)$ 分别代表 O_l 和 n_l 的概率

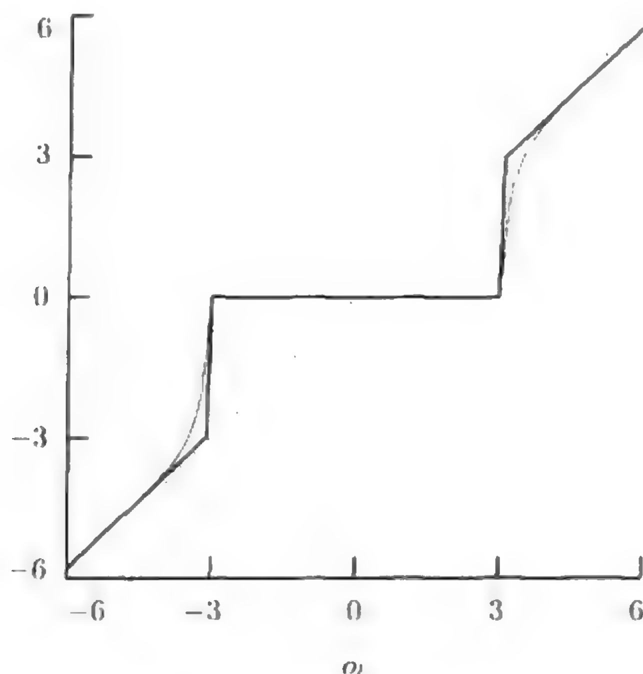


图 416 比较当 $\sigma^2_{G_t}=25$，$p_t=0.95$ 和 $\sigma^2_{n_t}=1$ 时的后验中值 $B_t(o_t)$（细曲线）和逼近条件中值 $U_t(o_t)$（粗曲线）. 点线表示的是对角线

密度函数. 现在，独立的随机变量的和的概率密度函数是他们密度的卷积（参见式(262a)），因此，O_t 的概率密度函数由 R_t 和 n_t 的概率密度函数的卷积给出：

$$f_{O_t}(o_t)=f_{R_t}*f_{n_t}(o_t)=p_tf_{R_t^{(1)}}*f_{n_t}(o_t)+(1-p_t)f_{R_t^{(2)}}*f_{n_t}(o_t).$$

既然 $R_t^{(1)}$ 和 n_t 是独立的有零均值的高斯随机变量，我们可以把 $f_{R_t^{(1)}}*f_{n_t}(\cdot)$ 解释为和 $O_t^{(1)}\equiv R_t^{(1)}+n_t$ 的概率密度函数 $f_{O_t^{(1)}}$. 如果我们使用两个均值为零的独立的高斯随机变量的和还是均值为零，方差是各自方差和的高斯随机变量的事实，则 $f_{O_t^{(1)}}(\cdot)$ 的概率密度函数可以看作是均值为零、方差为 $\gamma_t^2\sigma^2_{G_t}+\sigma^2_{n_t}$ 的高斯分布. 如果我们在式(256)中对于高斯概率密度函数设置为 $\mu=0$ 和 $\sigma^2=\gamma_t^2\sigma^2_{G_t}+\sigma^2_{n_t}$，并且与式(411a)中得到的 $A_t(\cdot)$ 的表达式相比较，能够得出 $A_t(o_t)=p_tf_{O_t^{(1)}}(o_t)$. 类似的论证告诉我们，$B_t(o_t)=(1-p_t)f_{O_t^{(2)}}(o_t)$，其中 $B_t(o_t)$ 也在式(411a)中给出，而且 $f_{O_t^{(2)}}(o_t)$ 是对于 $O_t^{(2)}\equiv R_t^{(2)}+n_t$ 的概率密度函数. 因此，给出 O_t 的实现 o_t，我们能把式(410f)的条件均值重新写成形式

$$E\{R_t\mid O_t=o_t\}=\frac{a_to_tp_tf_{O_t^{(1)}}(o_t)+b_to_t(1-p_t)f_{O_t^{(2)}}(o_t)}{p_tf_{O_t^{(1)}}(o_t)+(1-p_t)f_{O_t^{(2)}}(o_t)}. \tag{416}$$

比较式(411a)中 a_t 的定义和式(408)的简单收缩估计，我们看到，a_to_t 项相应于估计使得在 $O_t=o_t$ 处，$E\{(R_t^{(1)}-a_tO_t)^2\}$ 极小. 类似地，b_to_t 相应于估计在 $O_t=o_t$ 处，

使得 $E\{(R_l^{(2)}-b_lO_l)^2\}$ 极小. 因此条件均值可以看作是对信号分量 $R_l^{(1)}$ 和 $R_l^{(2)}$ 的最小均方误差估计，o_l 可能的权值来自于两个信号分量.

为了进一步理解，注意在最小均方估计中，我们需要 $E\{(R_l^{(1)}-a_lO_l)O_l\}=0$，如在练习[407]中所表示的，这表明随机变量 $R_l^{(1)}-a_lO_l$ 和 O_l 是不相关的，但是，由于它们在这里是联合高斯的，所以它们也是独立的. 因此

$$E\{(R_l^{(1)}-a_lO_l)\,|\,O_l\}=E\{R_l^{(1)}-a_lO_l\}=E\{R_l^{(1)}\}-a_lE\{O_l\}=0;$$

但是我们也有

$$E\{(R_l^{(1)}-a_lO_l)\,|\,O_l\}=E\{R_l^{(1)}\,|\,O_l\}-E\{a_lO_l\mid O_l\}=E\{R_l^{(1)}\,|\,O_l\}-a_lO_l.$$

给定 O_l 的实现 o_l，上面告诉我们 $E\{(R_l^{(1)}\mid O_l=o_l)\}=a_lo_l$，所以，非线性条件均值方法和线性最小二乘逼近给出联合高斯情况下相同的估计. 很明显对信号的第二个分量也是同样正确的，这就导致了对于为何这个模型的条件均值采用了式(416)的形式的深入洞察.

[2]混合高斯模型当然并不是式(410b)中统计信号分量的分布的唯一选择. 另一个可行的模型是广义高斯分布，其概率密度函数在式(257d)中给出，对此用标准数值积分技术容易求得条件均值(Walden, 1985).

10.5 独立同分布高斯小波系数

在 10.2 节和 10.4 节中我们已经看了阈值和收缩法则的设计，这基本上可以应用到任何规范正交变换系数的集合. 阈值或收缩在提取信号的优劣主要依赖于变换系数的性质. 我们已经在 10.1 节中说明离散小波变换可以简明地用时间频率性质的混合来表示信号. 现在我们集中看对小波系数应用阈值和收缩法则，并且在这样做的过程中我们讨论未知参数的估计. 当规范正交矩阵 $\mathcal{O}$ 用离散小波变换矩阵 $\mathcal{W}$ 代替，阈值或收缩实现时，我们谈论关于小波阈值或小波收缩.

本节我们看独立同分布高斯噪声，之后在 10.6 节考虑不相关非高斯噪声，然后在 10.7 节讨论相关的高斯噪声.

• 基于小波阈值

令考虑的模型为 $\mathbf{X}=\mathbf{D}+\boldsymbol{\varepsilon}$，其中 $\mathbf{D}$ 包含一个确定性信号. 应用离散小波变换 $\mathcal{W}$，我们得到

$$\mathbf{W}=\mathcal{W}\mathbf{D}+\mathcal{W}\boldsymbol{\varepsilon}=\mathbf{d}+\mathbf{e},$$

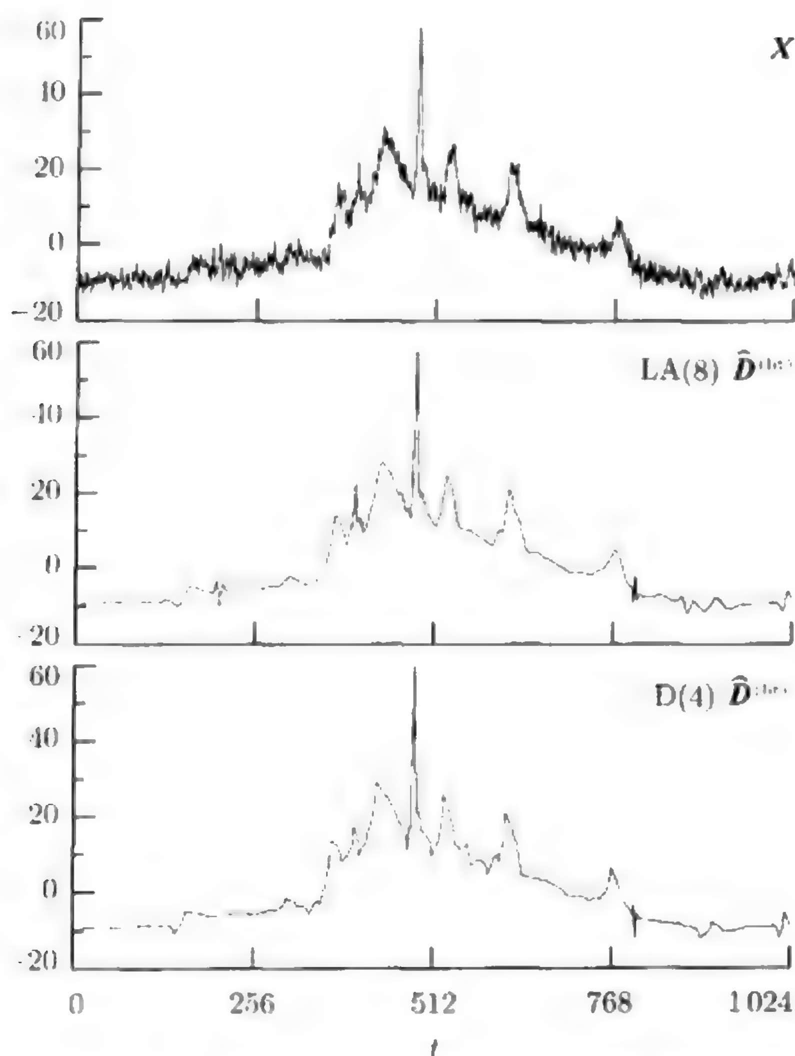


图 418 核磁共振(NMR)谱(最上面的图形)，以及基于小波硬阈值信号估计而用 $J_0=6$ 层的部分 LA(8)离散小波变换(中间的)和一个相似的 D(4)离散小波变换(最下面的)．在这两种情况下，我们用中位数绝对偏差标准离差估计$\hat{\sigma}_{(mad)}$决定噪声方差 σ_ϵ^2，这之后我们设置统一阈值水平$\hat{\delta}^{(u)}\equiv\sqrt{[2\hat{\sigma}_{(mad)}^2\log(N)]}$．这个核磁共振谱是从公共软件包 WaveLab 中提取的，它是由旧金山加利福尼亚大学放射系的 Andrew Maudsley 提供的(这些数据可以从这本书的网页中找到

我们认为这是式(398)的一个特殊情况．正如在 10.2 节中所讨论的，统一阈值适用于噪声是独立同分布高斯噪声，并有方差 σ_ϵ^2，使噪声变换系数 e 是独立同分布高斯的，还有共同的方差 σ_ϵ^2．Donoho and Johnstone(1994)要求用 J_0层部分离散小波变换而不是完全离散小波变换，其中 J_0 由使用者确定(参见 4.11 节)．在这个算法中，变换系数被包含在向量 $\mathbf{W}_1$，…，$\mathbf{W}_{J_0}$ 和 $\mathbf{V}_{J_0}$ 组成的 $\mathbf{W}$ 中，但是仅对在 $\mathbf{W}_k$ 向量中的系数取阈值；即 $\mathbf{V}_{J_0}$ 中的元素保持不变，因此 $\mathbf{X}$ 的部分归结于“粗”尺度 λ_{J_0} 自动分配于信号 $\mathbf{D}$.

阈值算法由下列的步骤组成.

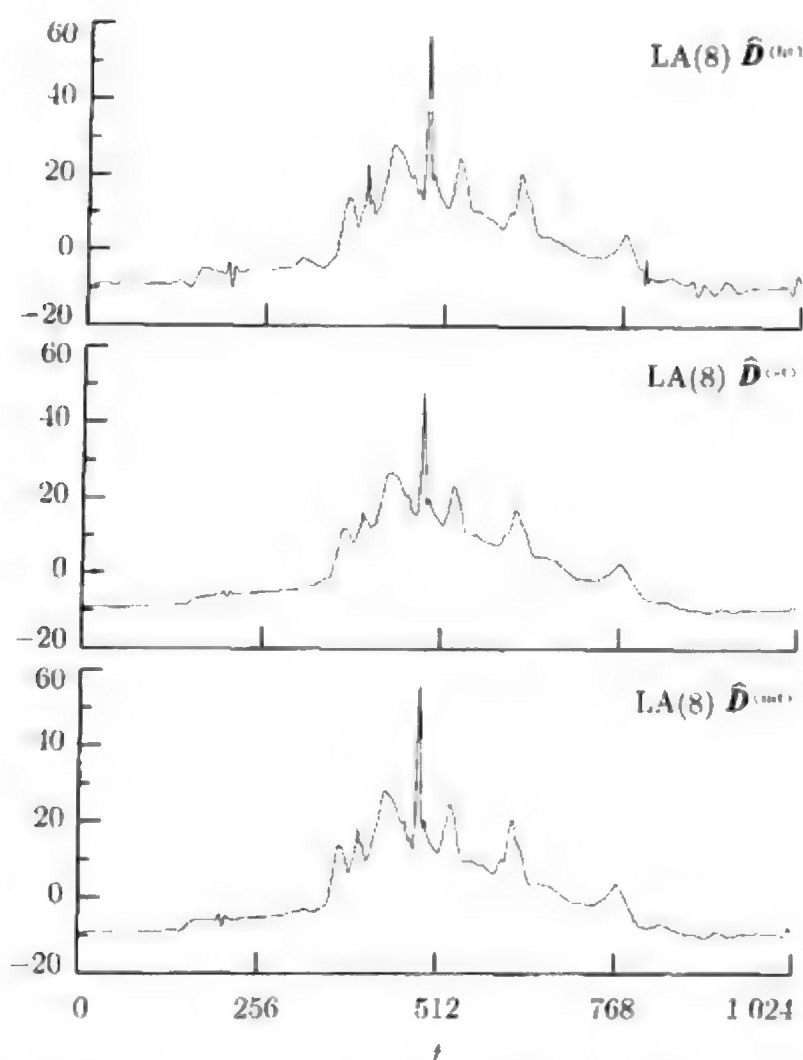


图 419　基于 $J_0=6$ 层的部分 LA(8)离散小波变换的核磁共振谱的阈值信号估计——从最上面到最下面——硬、软和中阈值(最上面的图形是图 418 的一个重复). 对这三个估计，我们用基于中位数绝对值偏差标准离差估计 $\hat{\sigma}_{(\mathrm{mad})}\doteq 1.645\ 38$ 的统一阈值水平 $\hat{\delta}^{(\mathrm{u})}\equiv\sqrt{[2\hat{\sigma}^2_{(\mathrm{mad})}\log(N)]}\doteq 6.126\ 22$

[1]计算一个 J_0 层部分离散小波变换来获取系数向量 $\mathbf{W}_1$，…，$\mathbf{W}_{J_0}$ 和 $\mathbf{V}_{J_0}$. 分量上说，我们有

$$W_{j,t}=d_{j,t}+e_{j,t}\quad j=1,\cdots,J_0;\ t=0,\cdots,N_j-1.$$

[2]其次确定阈值水平 $\delta=\delta^{(\mathrm{u})}$. 若 σ_ϵ^2 已知，用式(400d)计算 $\delta^{(\mathrm{u})}$. 如果 σ_ϵ^2 未知，用式(420)计算 $\hat{\sigma}_{(\mathrm{mad})}$，并用 $\hat{\delta}^{(\mathrm{u})}\equiv\sqrt{[2\hat{\sigma}^2_{(\mathrm{mad})}\log(N)]}$ 代替 $\delta^{(\mathrm{u})}$.

[3]对于 $W_{j,t}$，$j=1$，…，J_0 和 $t=0$，…，N_j-1，应用阈值法则之一(硬、中、软或固定的)，用 $W_{j,t}$ 代替 O_l 来获取阈值系数 $W_{j,t}^{(\mathrm{t})}$，之后把 $W_{j,t}^{(\mathrm{t})}$ 记为形式 $\mathbf{W}_j^{(\mathrm{t})}$，$j=1$，…，$J_0$.

[4]通过反变换 $\mathbf{W}_1^{(\mathrm{t})}$，…，$\mathbf{W}_{J_0}^{(\mathrm{t})}$ 和 $\mathbf{V}_{J_0}$ 得到 $\mathbf{D}$ 的估计 $\hat{\mathbf{D}}^{(\mathrm{t})}$.

如果σ_ϵ^2是未知的(通常——并不总是——在应用的情况下)，估计它在一个实际过程是建立在中位数绝对偏差(MAD)标准离差估计而使用在$\mathbf{W}_1$中$j=1$层$N/2$的系数. 由定义，标准离差估计为

$$\hat{\sigma}_{(\mathrm{mad})} \equiv \frac{\mathrm{median}\{|W_{1,0}|,|W_{1,1}|,\cdots,|W_{1,\frac{N}{2}-1}|\}}{0.6745}. \tag{420}$$

分母中因子0.674 5重新度量分子，所以$\hat{\sigma}_{(\mathrm{mad})}$也是高斯白噪声标准离差的合适估计. 对于从元素$\mathbf{W}_1$中计算出的启发式的理由是，除去可能的最大值最小尺度的小波系数应当由噪声决定. 中位数绝对偏差标准离差估计设计为稳健的最大变差，因此应当影响噪声方差而不是信号方差.

这种算法使用硬阈值的两个例子在图418中表示. 对于一个核磁共振谱(记录的谱有$N=1\,024$个数据，在上面的图形上部显示). 在这两种情况中，我们都用了中位数绝对偏差估计$\hat{\sigma}_{(\mathrm{mad})}$来设置阈值水平$\hat{\delta}^{(u)}$. 中间的图形表示当我们用LA(8)小波来计算$J_0=6$层的部分离散小波变换(从其中得到，$\hat{\sigma}_{(\mathrm{mad})} \doteq 1.645\,38$和$\hat{\delta}^{(u)} \doteq 6.126\,22$)，而底下的图形是用D(4)小波形成($\hat{\sigma}_{(\mathrm{mad})} \doteq 1.742\,45$，$\hat{\delta}^{(u)} \doteq 6.487\,66$)时得到的估计. 注意，两个估计都忠实地保留了在$t=500$附近尖峰的高度，更广的在核磁共振谱主要尖峰假设鲨鱼鳍出现在D(4)估计，这显然是由D(4)滤波器产生的假象(参见4.11节关于选择小波滤波器的讨论). 当我们使用光滑LA(8)滤波器的时候就没有这样的滤波假象.

图419比较了硬、软和中阈值估计. 核磁共振谱中的信号估计是基于在$J_0=6$层的部分LA(8)离散小波变换，阈值水平同样是通过中位数绝对偏差估计. 硬阈值估计(最上面的)是与图418中间的那个图相同. 注意有可注意的小尺度波动，特别在$t=200$和800附近. 软阈值估计(中)图形完全制止了这些波动，而且比硬阈值估计更光滑，不好的是有时抑制了尖峰. 特别是，尖峰的高度没有像在硬阈值情况下那样来忠实地处理. 更广泛的关于软阈值的缺点是能量层对于大系数在本质上改变了. 图419最底下的图表明中阈值减少了能量层的问题，而提供了足够抑制波动. 直觉上，最后一个估计比较好.

当列出的[1]～[4]的算法应用于软阈值时，Donoho and Johnstone(1994)把这个过程称为VisuShrink，但是注意我们可以把这归类于一个阈值法则，因为一些小波系数可用软阈值置为零. Donoho and Johnstone(1994)的定理1阐述了对VisuShrink的一个重要更正，即对所有可能的信号$\mathbf{D}$，用对数因子的理想风险$R(\hat{\mathbf{D}}^{(i)},\mathbf{D})$完成，这可能通过预言告诉我们哪个$W_{j,t}$的哪些系数被噪声支配：

$$R(\hat{\mathbf{D}}^{(st)},\mathbf{D}) \equiv E\{\|\hat{\mathbf{D}}^{(st)}-\mathbf{D}\|^2\} \leqslant [2\log(N)+1][\sigma_\epsilon^2+R(\hat{\mathbf{D}}^{(i)},\mathbf{D})].$$

代替使用统一阈值$\delta=\delta^{(u)}$，我们使用Stein无偏风险估计方法定义阈值水平. 对当前假设的具有方差σ_ϵ^2的独立同分布噪声小波系数，我们能使用$\hat{\sigma}^2_{(\mathrm{mad})}$估计

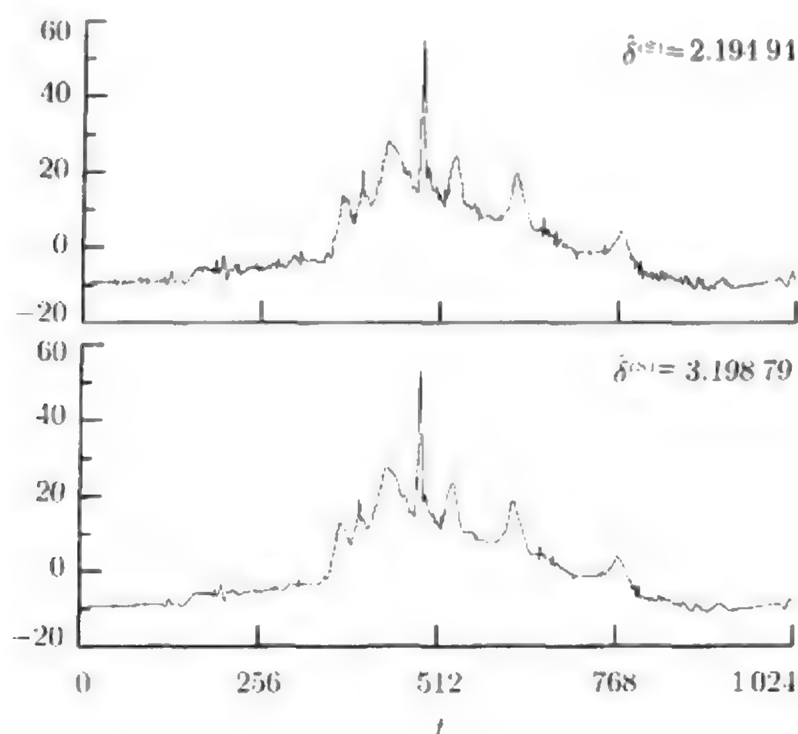


图 421 基于 $J_0=6$ 层的部分 LA(8)离散小波变换的核磁共振谱的软阈值和 Stein 无偏风险估计阈值水平 $\hat{\delta}^{(s)}$ 信号估计，这是分别用基于单位尺度的小波系数(最上面的图形)和所有六个尺度的小波系数的中位数绝对偏差尺度估计来计算的

σ_ϵ^2，并且给出 $W_{j,t}$ 的实现 $w_{j,t}$，之后可以通过设置它等于 δ 的平方根来确定 $\hat{\delta}^{(s)}$，δ^2 使下面式子最小：

$$\Upsilon(\delta^2) \equiv \sum_{j=1}^{J_0}\sum_{t=0}^{N_j-1}(2\,\hat{\sigma}^2_{(\mathrm{mad})} - w^2_{j,t} + \delta^2)1_{[\delta^2,\infty)}(w^2_{j,t}) \tag{421}$$

(参见(406a)). 在所有可能的使最小的值(即 $w^2_{j,t}$，$j=1, \cdots, J_0=6$，$t=0, \cdots, N_j-1$)用式(405b)给出的迭代来计算 $\Upsilon(\delta^2)$，我们对于核磁共振谱获得 $\hat{\delta}^{(s)} \approx 2.194\,94$，这小于相应的统一阈值水平 $\hat{\delta}^{(u)} \doteq 6.487\,66$. 由此得到的软阈值信号估计在图 421 的顶部而且比在图 418 和 419 中的估计有更多的噪声. 由于上面的 $\hat{\sigma}^2_{(\mathrm{mad})}$ 是所有小波系数中假设的共同方差的一个估计，我们能够构造另一个估计，它基于中位数绝对偏差中所有的系数而不是仅仅的单位尺度系数. 这个过程产生了噪声方差估计 2.052 70，阈值是 $\hat{\delta}^{(u)} \doteq 3.198\,79$，并且最小噪声信号估计在图 421 中的最下面给出.

最后，如 10.2 节所述，通过统一阈值或 Stein 的无偏风险估计的阈值水平的一个改变的来详细说明将用于两倍或者“漏一”的交叉赋值的方法，这两种都是用软阈值 Nason(1996)公式化小波系数. 对于两倍交叉赋值，开始分开

$\boldsymbol{X}\equiv[X_0, \cdots, X_{N-1}]^{\mathrm{T}}$为偶指标和奇指标下采样，即

$$\boldsymbol{X}^{(e)}\equiv[X_0,X_2,\cdots,X_{N-2}]^{\mathrm{T}} \quad 和 \quad \boldsymbol{X}^{(o)}\equiv[X_1,X_3,\cdots,X_{N-1}]^{\mathrm{T}}.$$

用阈值水平 δ，对两个下采样计算 J_0 层的部分离散小波变换，应用软阈值到小波系数，然后对阈值系数取逆离散小波变换，得到信号估计$\hat{\boldsymbol{D}}^{(e)}$和$\hat{\boldsymbol{D}}^{(o)}$. 其次，分别在$\hat{\boldsymbol{D}}^{(e)}$的元素$\hat{D}_t^{(e)}$ 和$\hat{\boldsymbol{D}}^{(o)}$的元素$\hat{D}_t^{(o)}$的基础上插值：

$$\hat{\boldsymbol{D}}^{(ie)}\equiv\frac{1}{2}\begin{bmatrix}\hat{D}_0^{(e)}+\hat{D}_1^{(e)}\\ \hat{D}_1^{(e)}+\hat{D}_2^{(e)}\\ \vdots\\ \hat{D}_{\frac{N}{2}-2}^{(e)}+\hat{D}_{\frac{N}{2}-1}^{(e)}\\ \hat{D}_{\frac{N}{2}-1}^{(e)}+\hat{D}_0^{(e)}\end{bmatrix} \quad 和 \quad \hat{\boldsymbol{D}}^{(io)}\equiv\frac{1}{2}\begin{bmatrix}\hat{D}_{\frac{N}{2}-1}^{(o)}+\hat{D}_0^{(o)}\\ \hat{D}_0^{(o)}+\hat{D}_1^{(o)}\\ \vdots\\ \hat{D}_{\frac{N}{2}-3}^{(o)}+\hat{D}_{\frac{N}{2}-2}^{(o)}\\ \hat{D}_{\frac{N}{2}-2}^{(o)}+\hat{D}_{\frac{N}{2}-1}^{(o)}\end{bmatrix}.$$

最后，分别用$\hat{\boldsymbol{D}}^{(ie)}$和$\hat{\boldsymbol{D}}^{(io)}$比较奇指标下采样和偶指标下采样，计算 δ 的范围，

$$\|\boldsymbol{X}^{(o)}-\hat{\boldsymbol{D}}^{(ie)}\|^2+\|\boldsymbol{X}^{(e)}-\hat{\boldsymbol{D}}^{(io)}\|^2. \tag{422}$$

使上面式子最小的 δ 表示为阈值水平$\hat{\delta}$. 因为这个层是建立在长度 $N/2$ 上的下采样. 而时间序列 $\boldsymbol{X}$ 长度为 N，Nason(1996)要求增长$\hat{\delta}$统一于要求统一阈值从 $N/2$增长到 N. 即我们一般用的阈值水平为

$$\hat{\delta}^{(\mathrm{tlcv})}\equiv\hat{\delta}\sqrt{\frac{\log(N)}{\log(N/2)}}.$$

为了应用"漏一"交叉赋值，确定一个阈值水平 δ，对 $1\leqslant l\leqslant N-2$ 重复下面的步骤.

[1]从 $\boldsymbol{X}=[X_0, \cdots, X_{N-1}]^{\mathrm{T}}$中除去 X_l，分割剩下的时间序列为"左"和"右"向量，即

$$\boldsymbol{X}^{(l)}\equiv[X_0,\cdots,X_{l-1}]^{\mathrm{T}} \quad 和 \quad \boldsymbol{X}^{(r)}\equiv[X_{l+1},\cdots,X_{N-1}]^{\mathrm{T}}.$$

[2]分别反映 $\boldsymbol{X}^{(l)}$和$\boldsymbol{X}^{(r)}$，扩展每个向量使它的长度是下一个 2 的指数. 分别折叠 X_{l-1}和 X_{l+1}获得列向量.

$$\boldsymbol{X}^{(le)}\equiv[X_{l-1},\cdots,X_{l-1},X_{l-1},\cdots,X_0,X_0,\cdots,X_{l-1}]^{\mathrm{T}} \text{ 和}$$

$$\boldsymbol{X}^{(re)}\equiv[X_{l+1},\cdots,X_{N-1},X_{N-1},\cdots,X_{l+1},X_{l+1},\cdots,X_{l+1}]^{\mathrm{T}}.$$

注意在 $\boldsymbol{X}^{(le)}$中项 N'的数是大于 $2l$ 的 2 的最小的指数，而 $\boldsymbol{X}^{(re)}$中的项是大于 $2(N-l-1)$的 2 的最小指数.

[3]对向量 $\boldsymbol{X}^{(le)}$和 $\boldsymbol{X}^{(re)}$运用完全的离散小波变换，然后对小波系数运用软阈值. 对阈值过的向量进行离散小波逆变换取得信号的估计$\hat{\boldsymbol{D}}^{(le)}$和$\hat{\boldsymbol{D}}^{(re)}$.

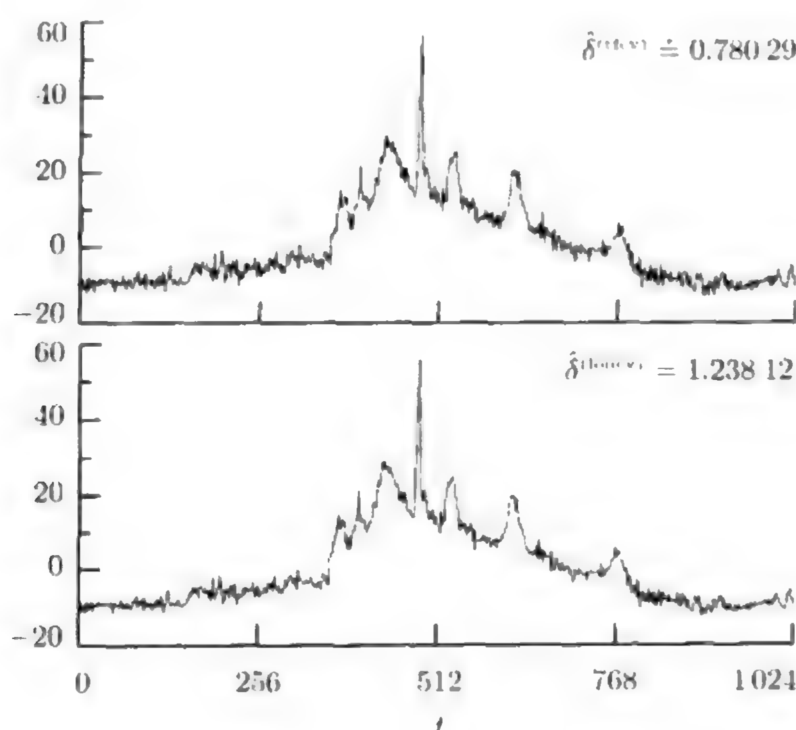


图 423　以层 $J_0=6$ 的部分 LA(8)离散小波变换的核磁共振谱的阈值信号估计，阈值为软阈值和两倍交叉赋值(最上面的图形)和"漏一"交叉赋值决定的阈值(最下面的图形)

[4]形成 $X_1^{(\delta)} \equiv (\hat{D}_{N'-1}^{(\text{le})}+\hat{D}_0^{(\text{re})})/2$，这是去掉 X_l，以 $\hat{\boldsymbol{D}}^{(\text{le})}$ 中的 $\hat{D}_{N'-1}^{(\text{le})}$ 为最下面点，$\hat{\boldsymbol{D}}^{(\text{re})}$ 中的 $\hat{D}_0^{(\text{re})}$ 为最上面点而得到的估计.

之后，能构造平方误差的和式 $\sum\limits_{l=1}^{N-2}(X_l^{(\delta)}-X_l)^2$. 这可对 δ 值的一个域进行计算，然后 $\hat{\delta}^{(\text{loocv})}$ 相应地通过极小化式(402)得到.

考虑用这三个交叉赋值算法之一选择 δ 有三个解释. 首先，当我们假设独立同分布高斯噪声 ε，实际上高斯性并不是必需的，但是 Nason(1996)注意到这个算法不那么好运行，如果我们在模型 $\boldsymbol{X}=\boldsymbol{D}+\boldsymbol{\varepsilon}$ 中的 ε 替换为相关噪声 η(这就是交叉赋值方法的著名问题). 第二，"漏一"交叉赋值扩展数据长度到 2 的指数，通过折叠 X_{l-1} 或 X_{l+1} 的方法保证了循环向量 $\boldsymbol{X}^{(\text{le})}$ 和 $\boldsymbol{X}^{(\text{re})}$ 输入到离散小波变换. 这种折叠的方法在处理硬过程中得到启发，不同于我们在 4.11 节中讨论的，其中我们赞成折叠取样中值，因为它保持取样方差的性质. 第三，"漏一"交叉赋值比两倍交叉赋值有优势，因为前者可对任意样本尺寸使用，然而后者更易于计算. 进而，后者用一个简单样本尺寸($N/2$)贯通始终(只是在最后产生 N 个值的时间序列，产生一个适合的项的一个适当的调整). "漏一"交叉赋值涉及了样本尺寸范围从 2 到至少 $2N-2$ 的一个混合，因此，不清楚得到的水平与样本尺寸 N 的关系.

作为例子，图 423 显示了用软阈值对核磁共振谱信号的估计，阈值水平有上面描述的两种交叉赋值方法. 这两种估计都是基于 LA(8)$J_0=6$ 层的部分离散小波变换. 对两倍和"漏一"交叉赋值估计水平是$\hat{\delta}^{(\mathrm{tfcv})}\doteq 0.780\,9$ 和$\hat{\delta}^{(\mathrm{loocv})}\doteq 1.238\,12$，这些相应于展示在图形的最上面和最下面的信号估计. 这两个估计是非常相似的，但是没有图 419 和图 421 中以统一阈值和 Stein 无偏风险估计阈值水平的估计表现的光滑. (关于小波阈值的交叉赋值的更进一步的讨论能在 Jansen 等，1997，Hurvich and Tsai，1998 中找到).

• 基于条件均值的小波收缩

现在，我们考虑的模型为 $\mathbf{X}=\mathbf{C}+\boldsymbol{\varepsilon}$，其中 $\mathbf{C}$ 为一个独立于 $\boldsymbol{\varepsilon}$ 的统计信号，应用$\mathcal{W}$，我们得到 $\mathbf{W}=\mathbf{R}+\mathbf{e}$，其中 $\mathbf{R}$ 包含统计信号的离散小波变换系数，假设独立于 $\mathbf{e}$. 伴随这种指定，我们可以应用条件均值收缩逼近. 收缩算法包含下列步骤.

[1]计算 J_0层部分离散小波变换获取 $\mathbf{W}_1$，…，$\mathbf{W}_{J_0}$ 和 $\mathbf{V}_{J_0}$. 对小波系数 $\mathbf{W}_j$ 的每一个分量，我们有

$$W_{j,t}=R_{j,t}+e_{j,t}\quad j=1,\cdots,J_0;\quad t=0,\cdots,N_j-1.$$

回忆我们的假设，信号系数 $R_{j,t}$有零均值；如 8.2 节所示，这就保证了小波系数提供了统计信号 $\mathbf{C}$ 是从一个 d 阶向后差分的统计过程是平稳的而来，并且，我们滤波器的长度取 L 使得 $L>2d$(注意没有理由假设 $\mathbf{V}_{J_0}$ 中的尺度系数有零均值，但这并没有关系，因为方法离开 $\mathbf{V}_{J_0}$ 不受影响).

[2]假设我们用稀疏信号模型(式(411b))而且假设参数不随层 j 或位置 t 而变化，即

$$R_{j,t}\stackrel{\mathrm{d}}{=}(1-\mathcal{I}_{j,t})\,\mathcal{N}(0,\sigma_G^2)\quad \text{和}\quad e_{j,t}\stackrel{\mathrm{d}}{=}\mathcal{N}(0,\sigma_\varepsilon^2),\tag{424}$$

其中对 $0\leqslant p\leqslant 1$，

$$\mathbf{P}[\mathcal{I}_{j,t}=1]=p\quad \text{并且}\quad \mathbf{P}[\mathcal{I}_{j,t}=0]=1-p.$$

因此信号和噪声分量对所有的 j 和 t 是同分布的. 如果我们令 σ_R^2 和 σ_W^2 分别为随机变量 $R_{j,t}$和 $W_{j,t}$的共同方差，我们有 $\sigma_R^2=(1-p)\sigma_G^2$ 和 $\sigma_W^2=\sigma_R^2+\sigma_\varepsilon^2$. 为了估计实现收缩法则需要的参数，我们能够

(a)取$\hat{\sigma}_\varepsilon=\hat{\sigma}_{(\mathrm{mad})}$；

(b)取平方系数 $W_{j,t}^2$，$j=1$，…J_0，$t=0$，…，N_j-1 的取样均值组成 $\hat{\sigma}_W^2$；并且

(c)对一个给出的 p 值，用$\hat{\sigma}_G^2=(\hat{\sigma}_W^2-\hat{\sigma}_\varepsilon^2)/(1-p)$估计$\hat{\sigma}_G^2$(典型的 p 本身被设为若干不同的值，并做出客观的评价来看，对某个特殊应用的最适用的——在图 425 中解释).

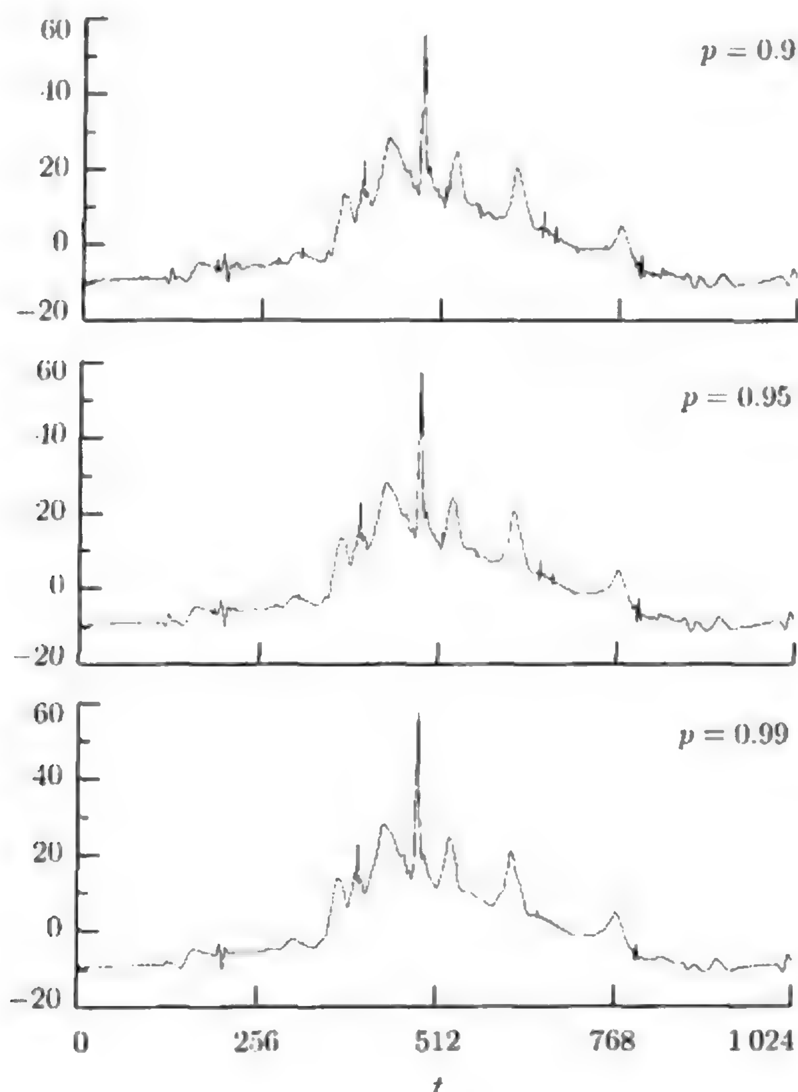


图 425 基于 $J_0=6$ 层的部分 LA(8) 小波变换和条件均值 $p=0.9$(最上面图形)、0.95(中间的)和 0.99(最下面)的核磁共振谱的收缩信号估计. 剩下的参数(即 σ_ϵ^2, σ_W^2 和 σ_G^2)的估计如在正文中解释的

[3]给出实现 $W_{j,t}$ 的 $w_{j,t}$ 之后，我们应用式(411c)来获得信号系数的收缩估计：

$$w_{j,t}^{(\mathrm{shr})} = \frac{\hat{b}}{1+\hat{c}} w_{j,t}, \tag{425}$$

其中对给定的 p 值，

$$\hat{b} = \frac{\hat{\sigma}_G^2}{\hat{\sigma}_G^2 + \hat{\sigma}_\epsilon^2}, \quad \hat{c} = \frac{p\sqrt{(\hat{\sigma}_G^2 + \hat{\sigma}_\epsilon^2)}}{(1-p)\,\hat{\sigma}_\epsilon} \mathrm{e}^{-w_{j,t}^2 \hat{b}/(2\hat{\sigma}_\epsilon^2)}.$$

[4]估计 $\boldsymbol{C}$ 的实现 $\boldsymbol{c}$，通过 $\hat{\boldsymbol{c}}^{(\mathrm{shr})}$ 用逆变换 $\boldsymbol{W}_1^{(\mathrm{shr})}$，…，$\boldsymbol{W}_{J_0}^{(\mathrm{shr})}$ 和 $\boldsymbol{V}_{J_0}$ 得到的，其中 $\boldsymbol{W}_j^{(\mathrm{shr})}$ 有分量 $w_{j,t}^{(\mathrm{shr})}$，$t=0$，…，N_j-1.

我们必须对在第[2]步中给出的参数估计的逼近做出两个解释. 首先，估计方差的减少一般是伴随着问题的；例如 $\hat{\sigma}_G^2$ 实际上能够是个负数. 第二，使用 $\hat{\sigma}_{(\mathrm{mad})}$

为对于σ_ϵ的一个估计取决于$j=1$层只含有很少信号的假定. 严格地讲，这与我们假设的随机变量$R_{j,t}$对所有的j和t是同分布的假设是不一致的，即$j=1$层小波系数与其他的层有相同的信号结构——这就是估计$\hat{\sigma}_{(\mathrm{mad})}$应当没有理由更喜欢这些小波系数而不是其他的关于估计σ_ϵ. 因此我们应当修改$\hat{\sigma}_{(\mathrm{mad})}$的定义使应用所有的小波系数. 有没有这个修改，$\hat{\sigma}_{(\mathrm{mad})}$的稳健性质使它成为$\sigma_\epsilon$在相当少的信号项出现时为一个好的估计(如$p$接近于1的时候，例如0.95或0.99也是这样的).

图425显示了对于值p为0.9，0.95，0.99应用于核磁共振谱的逼近. 对于$p=0.99$的结果与用统一阈值的硬阈值法结果是相同的(图419的最上面的图形).

• *基于贝叶斯方法小波收缩*

作为用贝叶斯逼近的解释，我们现在返回到Vidakovic(1998)的算法. 在10.4节，我们列出了对每个变换系数的分离模型下的逼近. 然而，每个模型依赖于三个参数，即变换噪声系数先验分布的超参数v_l、变换信号系数t分布的两个参数ϑ_l和κ_l. 如果这三个参数已知并且我们假设，例如平方误差损失，我们可以构造式(414b)的贝叶斯收缩法则来估计基于相应的观察变换系数的每个信号系数，因此我们不必任何严格假设变换系数的联合概率密度函数(参见10.7节对参数已经知道的信号估计问题)；另一方面，如果参数未知，我们面临对每个变换系数估计三个参数. 为了简化估计问题，我们必须对这个模型加以限制.

在离散小波变换情况下，Vidakovic(1998)假设小波系数$W_{j,t}$都依赖于三个相同参数，即v(在噪声小波系数方差是指数先验分布的普通超参数)及ϑ和κ(对于信号小波系数的概率密度函数的普通参数). 额外地，他假设每一个噪声系数是关于一个简单的随机变量σ_ϵ^2是有条件的. 因此，条件加于这个随机变量，噪声离散小波变换系数$\mathbf{e}=\mathcal{W}\boldsymbol{\epsilon}$实际上是独立同分布高斯的(Chipman等，1997和Abramovich等，1998，从不同的角度考虑都得到这个噪声模型). 对J_0层部分离散小波变换的每个分量，Vidakovic的逼近用

$$W_{j,t}=R_{j,t}+e_{j,t}.$$

模型则是

$$e_{j,t}\,|\,\sigma_\epsilon^2 \stackrel{\mathrm{d}}{=} \mathcal{N}(0,\sigma_\epsilon^2),\quad \text{因此} W_{j,t}\,|\,R_{j,t},\sigma_\epsilon^2 \stackrel{\mathrm{d}}{=} \mathcal{N}(R_{j,t},\sigma_\epsilon^2),$$

其中σ_ϵ^2有一个先验(指数)分布(式(413c)). 给出观察的小波系数，这个模型有三个未知参数(v，ϑ，κ)和未知的随机变量σ_ϵ^2的实现. 由于对所有的j和t，信号和噪声项是同分布的，我们可以重新让σ_R^2，σ_ϵ^2和σ_W^2分别表示随机变量$R_{j,t}$，$e_{j,t}$和$W_{j,t}$的共同方差. 因此，我们能够获得下面的参数.

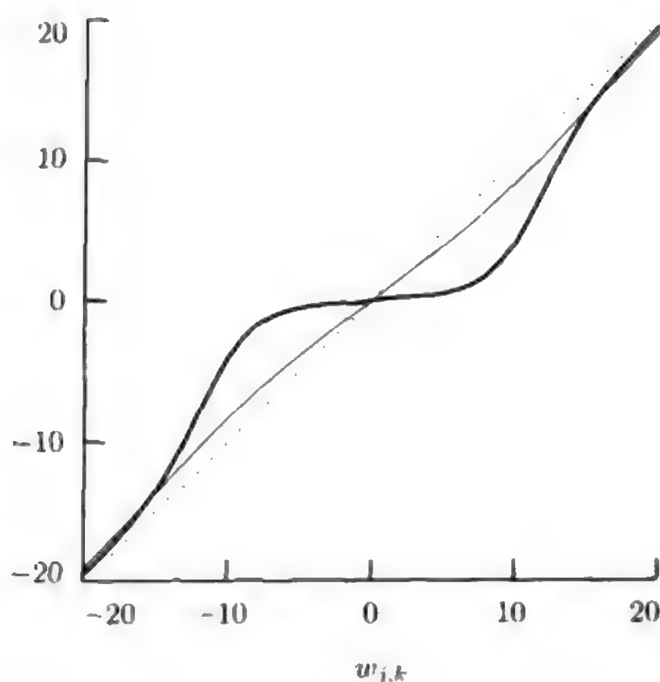


图 427 对于 Vidakovic(1998)对算法应用于核磁共振谱作为 $w_{j,k}$ 函数的贝叶斯法则 $B_2(w_{j,k})$. 细的和粗曲线描绘了假设自由度 ϑ 分别是 5 和 2.01. 点线标记的是对角线. 相应的信号估计在图 428 中给出

[1]我们能用$\hat{\sigma}_{(\mathrm{mad})}$(式(420)估计 σ_ϵ 的实现，由于在这个实现上的条件，独立同分布的假设是适当的.

[2]对超参数 v，我们知道 $E\{\sigma_\epsilon^2\}=1/v$，因为 σ_ϵ^2 有指数分布. 因此 Vidakovic 建议取 $\hat{v}=1/\hat{\sigma}_\epsilon^2$.

[3]由于零均值的随机变量 $R_{j,t}$ 有 ϑ 自由度的 t 分布和尺度参数 κ. 我们知道它们的共同方差是

$$\sigma_R^2 = \frac{\kappa^2 \vartheta}{\vartheta - 2}.$$

对于 κ 的另一个值能够特别取为

$$\hat{\kappa} = \left(\hat{\sigma}_R^2 \frac{\vartheta - 2}{\vartheta}\right)^{1/2},$$

其中$\hat{\sigma}_R^2$ 是从$\hat{\sigma}_R^2 = \hat{\sigma}_W^2 - \hat{\sigma}_\epsilon^2$ 求得的，而$\hat{\sigma}_W^2$ 是由 $W_{j,t}^2$ 在 $j=1, \cdots, J_0$，$t=0, \cdots, N_j-1$ 上的平均形成的；然而，正如前面提出的估计方差的削减是成问题的.

[4]Vidakovic(1998)发现置自由度 ϑ 的值为 5 都是很满意的.

我们注意到，不像 Vidakovic(1998)，Chipman 等(1997)和 Abramovich 等(1998)也允许超参数随层 j 变化.

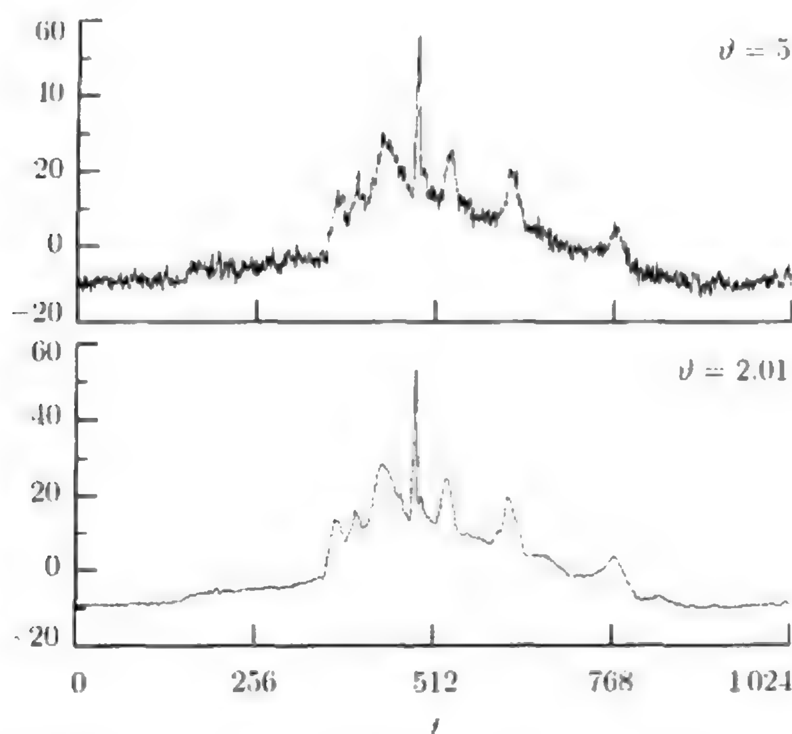


图 428 基于 $J_0=6$ 层的部分 LA(8) 小波变换和在式(414b)中给出的 Vidakovic(1998)公式化的贝叶斯法则 $B_2(\cdot)$(对每个估计的特别法则在图 427 中给出了)的核磁共振谱的收缩信号估计. 这两个估计的不同之处在于选择信号概率密度函数的自由度, 剩下的两个参数估计如在正文中所描写的

作为例子, 让我们对核磁共振数据应用 Vidakovic 过程, 再次使用 $J_0=6$ 层 LA(8)部分离散小波变换. 为了估计所需要的参数 υ, 我们计算$\hat{\sigma}_{(\mathrm{mad})}\doteq 1.645\ 38$, 产生$\hat{\upsilon}\doteq 0.369\ 38$; 为了估计 κ, 我们计算$\hat{\sigma}_W^2\doteq 14.367\ 55$ 和$\hat{\sigma}_R^2\doteq 11.660\ 29$. 并且置 $\vartheta=5$, 从其中我们得到$\hat{\kappa}\doteq 2.645\ 03$. 由此而来的贝叶斯法则 $B_2(\cdot)$可用式(414b)的数值积分来计算, 并且在图 427 中用细曲线绘出. 相应的估计信号在图 428 的最上面. 这里贝叶斯法则展示了少量的收缩, 所以信号估计和原始数据(图(418)最上面的图形)是相当类似的. 然而, 我们假设变换信号概率密度函数 $f_R(\cdot)$是具有设置自由度 $\vartheta=5$ 所展示的图形有更重的尾巴, 即我们允许少量的信号小波系数为很大的这种可能性, 如这种可能对给出观察核磁共振尖峰的理由. 如果我们使 $\mathrm{var}\{R\}$确定在它的估计$\hat{\sigma}_R^2\approx 11.660\ 29$, 我们可使 $f_R(\cdot)$的尾巴由 ϑ 减小到趋向 2 的程度来决定(这样也有收缩关于零对称区间的效果, 捕捉一个给出 95%概率的变换信号系数). 客观选择 $\vartheta=2.01$ 产生贝叶斯法则在图 427 中表示为粗曲线, 估计表示在图 428 最下面的图形. 信号估计现在更多地保持那些获取的条件均值估计得到的(图 425), 虽然前者表现非常小的波形而后者更不清楚.

10.5 节的评论与扩展

[1]如第 5 章所述，离散小波变换的一个隐藏问题在于我们在哪里"插入"时间序列它是敏感的. Coifman and Donoho(1995)，Lang 等(1996)和 Bruce 等(1999)找到了寻找减轻这种敏感性的方法，即用所谓的"循环旋转"(我们在练习[5.1]中简单提到了这个概念). 当对一个长度 N 为 2^{J_0} 的整数倍的时间序列，使用了 J_0 层的部分离散小波变换时，基本的思想是不仅对原始序列 $\mathbf{X}$，而且对应用所有可能的感兴趣的循环平移序列，即 $\mathcal{T}^n\mathbf{X}$ 应用一个信号提取过程. 为简单化，我们主要看一个基于小波的阈值算法应用到模型 $\mathbf{X}=\mathbf{D}+\boldsymbol{\epsilon}$，其中 $\mathbf{D}$ 是一个确定性信号，$\boldsymbol{\epsilon}$ 是时独立同分布高斯随机变量，每一个方差均为 σ_ϵ^2. 如果 $\mathcal{T}^n\mathbf{X}$ 导出估计 $\hat{\mathbf{D}}_n^{(t)}$，则我们可获得一个很少受到开始时间影响的估计，这是通过平均这些单个的估计，在每一次平移之后，使它同原始序列 $\mathbf{X}$ 一致，即"循环旋转"产生信号估计

$$\widetilde{\mathbf{D}}^{(t)} \equiv \frac{1}{2^{J_0}}\sum_{n=0}^{2^{J_0}-1}\mathcal{T}^{-n}\hat{\mathbf{D}}_n^{(t)}$$

(我们不需要考虑 $n\geq 2^{J_0}$：这样的平移产生一个 $\mathcal{T}^n\mathbf{X}$，其离散小波变换是冗余的，因为 $\mathcal{T}^{-n}\hat{\mathbf{D}}_n^{(t)}=\mathcal{T}^{-(n \bmod 2^{J_0})}\hat{\mathbf{D}}_{n \bmod 2^{J_0}}^{(t)}$). 在上面参考文献里面列出的模拟实验表明在平方均值误差意义下 $\widetilde{\mathbf{D}}^{(t)}$ 优于 $\hat{\mathbf{D}}_0^{(t)}$.

因为对各种各样的 $\mathcal{T}^n\mathbf{X}$，离散小波变换可以从 $\mathbf{X}$ 的极大重叠离散小波变换中提取(在适当的再规范化后)，"循环旋转"实际上可以有效地借助于极大重叠离散小波变换，如下(为了简单，我们假设，在相应的离散小波变换中，我们使用式(400d)的统一阈值水平即 $\delta^{(u)}\equiv\sqrt{2\sigma_\epsilon^2\log(N)}$，而 σ_ϵ^2 已知). 首先我们计算 J_0 层极大重叠离散小波变换来获得系数向量 $\widetilde{\mathbf{W}}_j$，$j=1, \cdots, J_0$ 和 $\widetilde{\mathbf{V}}_{J_0}$. 之后，我们对 $\widetilde{\mathbf{W}}_j$ 的每个元素用阈值法则之一，使用水平依赖的阈值 $\tilde{\delta}_j^{(u)}\equiv\sqrt{[2\sigma_\epsilon^2\log(N)]/2^j}$(注意，其中 $\tilde{\delta}_j^{(u)}=\delta^{(u)}/2^{j/2}$ 模拟了在 5.4 节中给出的第 j 层的离散小波变换和极大重叠离散小波变换的关系，即 $\tilde{h}_{j,l}=h_{j,l}/2^{j/2}$). 令 $\widetilde{\mathbf{W}}_j^{(t)}$ 表示 $\widetilde{\mathbf{W}}_j$ 阈值项，我们可以用 $\widetilde{\mathbf{W}}_j^{(t)}$，$j=1, \cdots, J_0$ 和 $\widetilde{\mathbf{V}}_{J_0}$ 的极大重叠离散小波逆变换来取得 $\widetilde{\mathbf{D}}^{(t)}$.

两个注释是适用的. 首先，循环旋转后的论证，假设 $\mathbf{X}$ 的样本尺寸 N 是 2^{J_0} 的整数倍. 我们实际上可以对任意取样尺度应用上面的过程，因此引出的对总体 N 有效的 $\widetilde{\mathbf{D}}^{(t)}$ 的定义. 第二，如果 σ_ϵ^2 未知，我们采用中位数绝对偏差尺度估计而用极大重叠离散小波变换来做. 为此，我们定义基于极大重叠离散小波变换，标准离差的中位数绝对偏差估计为

$$\tilde{\sigma}_{(\mathrm{mad})} \equiv \frac{2^{1/2}\mathrm{median}\{|\widetilde{W}_{1,0}|, |\widetilde{W}_{1,1}|, \cdots, |\widetilde{W}_{1,N-1}|\}}{0.6745}.$$

可以与上面的式(420)比较，基于离散小波变换和基于极大重叠离散小波变换的估计的主要区别是在上式分子上的额外的"$2^{1/2}$"，这是保持在单位尺度离散小波变换和极大重叠离散小波变换系数之间的关系，即 $W_{1,t}=2^{1/2}\widetilde{W}_{1,2t+1}$.

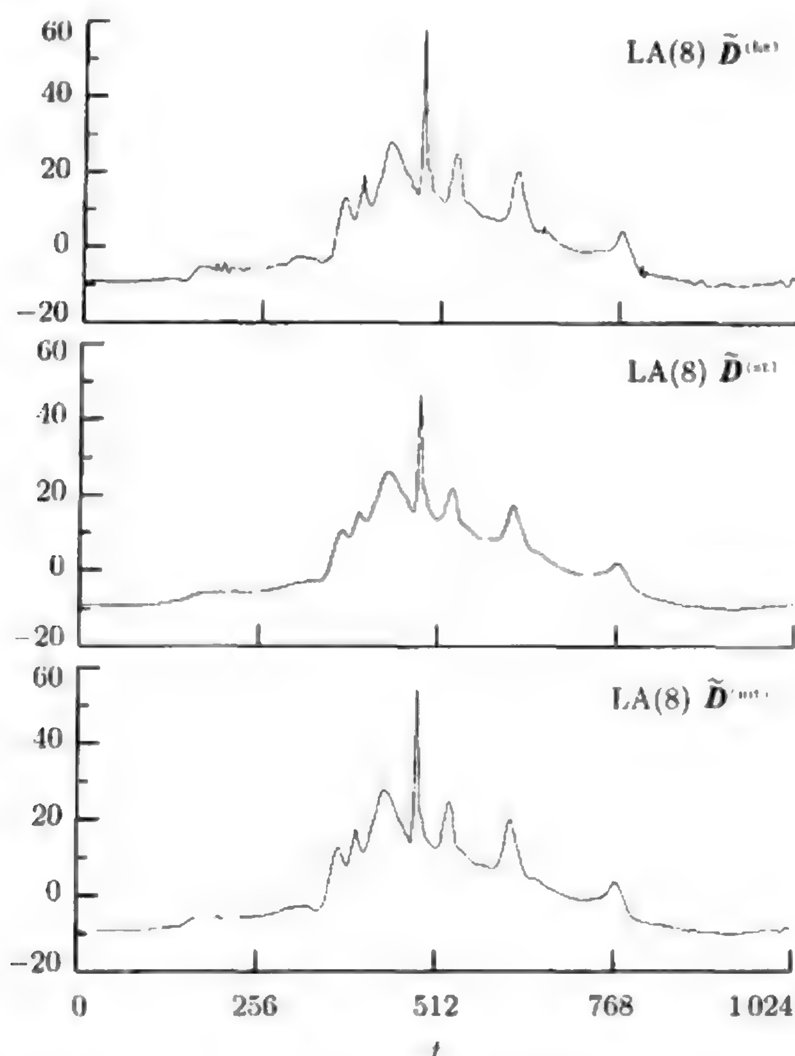


图 430 基于 $J_0=6$ 层的部分 LA(8)极大重叠离散小波变换核磁共振谱的阈值信号估计—从最上面到最下面—硬、软和中阈值(图 419 有对离散小波变换相应的图形). 每一个估计用统一阈值水平 $\tilde{\delta}_j^{(u)}\equiv\sqrt{[2\tilde{\sigma}^2_{(mad)}\log(N)/2^j]}\doteq 6.49673/2^{j/2}$ 是通过极大重叠离散小波变换的中位数绝对偏差标准离差估计 $\tilde{\sigma}_{(mad)}\doteq 1.74489$ 来计算的

作为例子，图 430 表示了在核磁共振谱的信号基于 $J_0=6$ 层的 LA(8)极大重叠离散小波变换的硬、软及中阈值估计. 每一个估计依赖层统一阈值 $\tilde{\delta}_j^{(u)}\equiv\sqrt{[2\tilde{\sigma}^2_{(mad)}\log(N)/2^j]}\doteq 6.49673/2^{j/2}$，对此我们算出 $\tilde{\sigma}_{(mad)}\doteq 1.74489$. 与图 419 中相应的基于离散小波变换的估计相比，小尺度的波纹在硬阈值估计的 $t=200$，800附近明显比在基于极大重叠离散小波变换的估计稀薄，其中峰值在两个估计中是大体可比的. 基于极大重叠离散小波变换的软阈值和中阈值比他们相应的离散小波变换光滑. 基于极大重叠离散小波变换估计中主峰的高度是低于基于离散小波变换的估计，这可能是后者的一个优点(在偏方差中，额外平均隐含在基于极大重叠离散小波变换的估计来降低方差引起的更大的偏差).

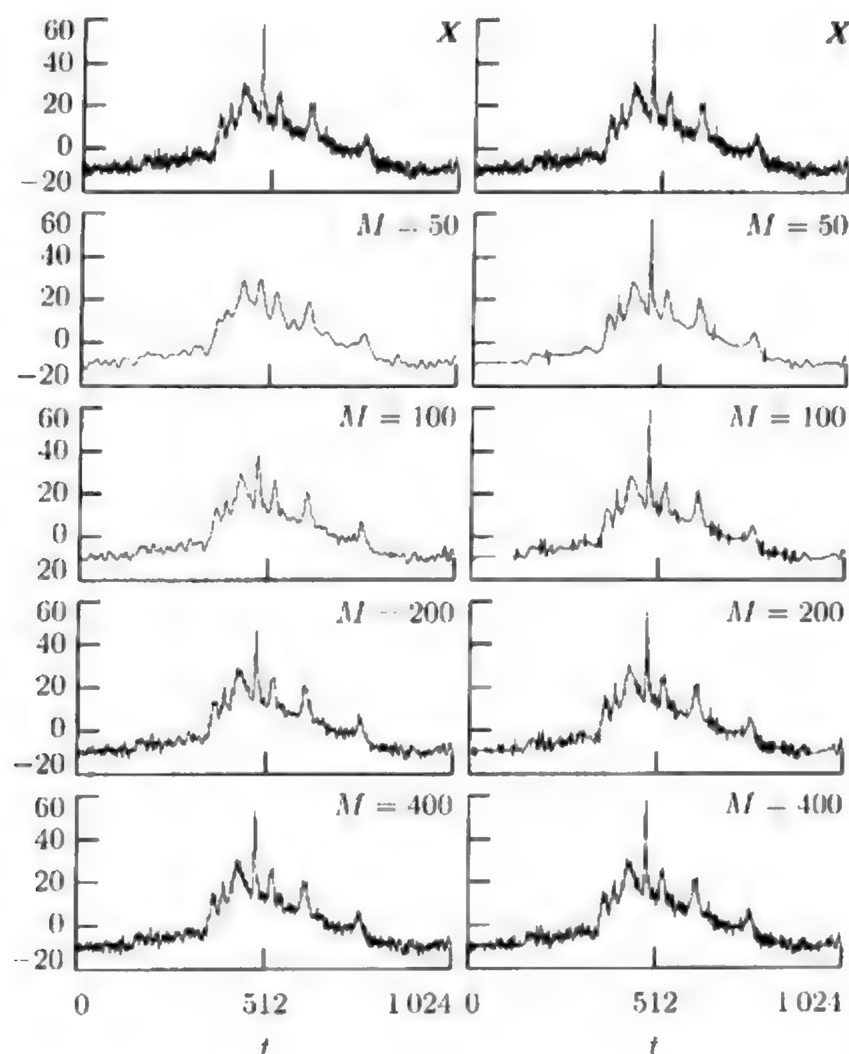


图 431 基于保持 M 个幅值最大的系数在正交离散傅里叶变换中(左列)和 $J_0=6$ 层部分 LA(8)离散小波变换(右列)用硬阈值方法的核磁共振谱(最上面的图形)的去噪. $M=$ 50, 100, 200 和 400(分别为第二行到第五行)

[2]用核磁共振谱 $\boldsymbol{X}$ 进一步比较在时间序列中哪种性质被正交离散傅里叶变换和离散小波变换很好的保持的方法是很有益的.(这将在 10.1 节中讨论). 在图 431 中做出了这种比较. 第一行展示是 $\boldsymbol{X}$ 本身的图, 下面的四行表示在 $\boldsymbol{X}$ 中的信号的估计用 $M=50$, 100, 200 和 400 个最大幅值的正交离散傅里叶变换系数(左列)和 $J_0=6$ 层部分 LA(8)离散小波变换系数(右列). 特别注意到, 在 $\boldsymbol{X}$ 中 $t=500$ 附近的尖峰, 这仅用 50 个离散小波变换系数就可以忠实地获取, 而它本质上是在 $M=50$ 正交离散傅里叶变换综合的时候丢失的, 而仍然可见在 $M=$ 400 综合中明显低估了. 对核磁共振谱这样的数据, 保持这样的峰值是很重要的, 并且离散小波变换可以把峰值很好地孤立, 在几个系数中, 而正交离散傅里叶变换则把它们传到了许多系数中(一个时间域的尖峰在频域覆盖开来, 反之亦然). 另一方面, $M=50$ 的正交离散傅里叶变换综合在时间序列的开始和结束获

取某些高频震荡，而在相应的离散小波变换综合中遗失了——核磁共振分析者们大概认为是震荡噪声的一部分而不是信号的一部分. 因此更喜欢离散小波变换的这一方面. 我们可以想像其他的时间序列，然而，在高频震荡是信号很重要的一部分，尖峰将是不希望出现的—对这样的序列的分析将找到正交离散傅里叶变换而不是离散小波变换来提取这种信号. 其中很重要的信号是考虑对特别问题的细节，即将主宰的是哪种变换对提取特别的信号最有用. 离散小波变换适用于广泛类型的信号，但是在这节中我们用它来估计信号的方法可能有问题，例如，已经列出来的(参见 Bruce 等 1994，Bruce and Gao，1996a，对一个“光滑清除”的小波变换强烈反对).

10.6 不相关非高斯小波系数

在 10.5 节中，我们见的小波阈值或收缩是在独立同分布高斯噪声 ε 的假设之下，这给出了独立同分布的高斯噪声的离散小波变换系数 $\mathbf{e}=\mathcal{W}\boldsymbol{\varepsilon}$. 本节我们看到，附加的噪声仍然是独立同分布但是非高斯的情况. 这就意味着，噪声离散小波变换系数 $\mathbf{e}$ 一般不是同分布的，但它们仍然是不相关的. 感兴趣的特别例子是通过周期图小波阈值谱估计的方法.

• *谱估计算法*

假设我们有一个时间序列可认为是零均值的稳定过程 X_0，…，X_{N-1} 的实现，其谱密度函数为 $S_X(\cdot)$，取样间隔为 Δt. 假设 N 是 2 的幂. 对序列 $\{X_0, \cdots, X_{N-1}\}$ 使用标准快速傅里叶变换算法. 在傅里叶频率点 $f_k=k/(N\Delta t)$，$k=0, \cdots, M$，其中 $M=N/2$(参见式(271a))，我们能用式(269a)来计算

$$Y^{(p)}(f_k)=\log(\hat{S}_X^{(p)}(f_k))+\gamma=\log(S_X(f_k))+\varepsilon(f_k).$$

上面说的是对数周期化——加了欧拉(Euler)常数 $\gamma\doteq 0.577\,21$ 以后——能写作一个信号(真对数谱密度函数)加了一个零均值的噪声. 而对 $o<f_k<f_N$ 的 $\{\varepsilon(f_k)\}$ 组成了一个逼近独立同分布的随机变量的集合，使

$$\varepsilon(f_k)\stackrel{d}{=}\log(\chi_2^2)+\gamma-\log(2)$$

(参见式(270c)). 上面推出 $\sigma_\varepsilon^2\equiv\operatorname{var}\{\varepsilon(f_k)\}=\pi^2/6$. 小波阈值算法能怎样使用于这个信号估计问题呢?

用小波阈值来产生一个对稳定过程的对数谱密度函数的光滑估计的基本算法是很简单的，如 Moulin(1994)或 Gao(1993，1997)的实现，由下面四个基本步骤组成：

[1]计算在傅里叶频率的周期图的对数，组成 $Y^{(p)}(f_k)=\log(\hat{S}_X^{(p)}(f_k))+\gamma$.

[2]对 $\{Y^{(p)}(f_k)\}$ 使用 J_0 水平的部分离散小波变换.

[3]对于实验的小波系数，使用阈值过程，产生阈值后的系数集合.

[4]对阈值后的实验小波系数使用离散小波逆变换，产生对数谱密度函数在傅里叶频率点的光滑估计.

下面让我们一步一步仔细地看一下这个算法.

[1]如果我们想要把对数谱密度函数估计放在独立同分布附加噪声模型上，则必须严格用 $Y^{(p)}(f_k)$使得 $0<f_k<f_N$ 即 $k=0,\cdots,M-1$(随机变量 $Y^{(p)}(f_0)$和 $Y^{(p)}(f_M)$)有相关于 $\log(\chi_1^2)$随机变量的分布而不是 $\log(\chi_2^2)$随机变量的分布. 不幸的是，这产生了一个样本尺寸为 $M-1$，而不是 2 的幂，因此不符合通常对离散小波变换的限制. 相应地，我们考虑包括获得 $Y^{(p)}(f_0)$的项，用向量记号

$$\begin{bmatrix} Y^{(p)}(f_0) \\ \vdots \\ Y^{(p)}(f_{M-1}) \end{bmatrix} = \begin{bmatrix} \log(S_X(f_0)) \\ \vdots \\ \log(S_X(f_{M-1})) \end{bmatrix} + \begin{bmatrix} \varepsilon(f_0) \\ \vdots \\ \varepsilon(f_{M-1}) \end{bmatrix}, \tag{433a}$$

这给我们一个可观察的向量，它的长度在包括了一个不规则分布的简单项(它的影响对于大的 $M=N/2$ 可忽略)为 2 的幂. 上面的公式并不能完全满足，因为离散小波变换使用循环滤波. 由于不要求 $S_X(0)$和 $S_X(f_N)$彼此任意近，我们能够在频率接近于零和 f_N 的扭曲估计$\log(S_X(f))$来结束. 一种减轻这种扭曲的自然的方法是通过定义一个周期为 $2f_N$ 的偶周期函数，恢复对实值过程的谱密度函数. 相应地，让我们使用离散小波变换对式(433a)展开变形，即

$$\mathbf{Y}^{(p)} \equiv \begin{bmatrix} Y^{(p)}(f_0) \\ \vdots \\ Y^{(p)}(f_{2M-1}) \end{bmatrix} = \begin{bmatrix} \log(S_X(f_0)) \\ \vdots \\ \log(S_X(f_{2M-1})) \end{bmatrix} + \begin{bmatrix} \varepsilon(f_0) \\ \vdots \\ \varepsilon(f_{2M-1}) \end{bmatrix} \equiv \mathbf{D} + \boldsymbol{\varepsilon}. \tag{433b}$$

由于 $Y^{(p)}(f_{2M-k}) = Y^{(p)}(f_k)$对 $k=1,\cdots,M-1$ 成立，上面是非常接近于我们在式(433a)左边需要的分析的，而使用式(140)反射边界条件算法(我们两次用不规则的随机变量 $Y^{(p)}(f_0)$和 $Y^{(p)}(f_M)$而不是用 $Y^{(p)}(f_0)$偏离这个算法). Moulin(1994)做了一个类似的论证，实际上，形成谱估计的一个标准方法是在一个完全周期中光滑周期图循环(作为例子参见式(271c)).

[2]计算 $\mathbf{Y}^{(p)}$ 的一个 J_0 层的部分离散小波变换分解以得到向量 $\mathbf{W}_1^{(p)},\cdots,\mathbf{W}_{J_0}^{(p)}$ 和 $\mathbf{V}_{J_0}^{(p)}$. 逐个分量是

$$W_{j,t}^{(p)} = d_{j,t} + e_{j,t} \quad j=1,\cdots,J_0;\ t=0,\cdots,2M_j-1, \tag{433c}$$

其中 $M_j \equiv M/2^j = N/2^{j+1}$. 这里$\{d_{j,t}\}$和$\{e_{j,t}\}$分别是信号 $\mathbf{D}$ 和噪声 $\boldsymbol{\varepsilon}$ 的第 j 层小波系数.

[3]对 $\mathbf{W}_1^{(p)},\cdots,\mathbf{W}_{J_0}^{(p)}$ 使用阈值算法得到 $\mathbf{W}_1^{(t)},\cdots,\mathbf{W}_{J_0}^{(t)}$. 注意即使离散小波变换使用到一个长度为 $2M$ 的向量，我们必须计算适合样本尺寸 M 的阈值水平

(在式(433a)中独立的噪声项是 M 个，并且在 $\boldsymbol{Y}^{(p)}$ 中 M 个附加项来源于在合理意义下处理边界条件).

[4]用 $\boldsymbol{W}_1^{(p)}$，…，$\boldsymbol{W}_{J_0}^{(p)}$ 和 $\boldsymbol{V}_{J_0}^{(p)}$ 的逆变换得到 $\boldsymbol{D}$ 的估计 $\hat{\boldsymbol{D}}^{(t)}$. 令 $\hat{\boldsymbol{D}}_k^{(t)}$ 表示 $\hat{\boldsymbol{D}}^{(t)}$ 的第 k 个元素，我们现在能够通过下式形成谱估计：

$$\hat{S}_X^{(pt)}(f_k)=\begin{cases}\mathrm{e}^{\hat{D}_0^{(t)}}, & k=0;\\ \dfrac{1}{2}\left(\mathrm{e}^{\hat{D}_k^{(t)}}+\mathrm{e}^{\hat{D}_{2M-k}^{(t)}}\right), & k=1,\cdots,M-1;\\ \mathrm{e}^{\hat{D}_M^{(t)}}, & k=M.\end{cases}$$

我们注意到，与典型的谱密度函数估计不同，阈值估计 $\exp(\hat{\boldsymbol{D}}_k^{(t)})$ 不需要关于 $f=f_{\mathcal{N}}$ 对称. 这就是我们选择在频率 f_k 和 f_{2M-k} 平均估计来组成在第[4]步中的 $\hat{S}_X^{(pt)}(f_k)$ 的原因(这种对称性的源头是练习[10.8]的目标).

• **系数的统计性质和阈值规范**

由于离散小波变换的局部性质，因为 $\boldsymbol{Y}^{(p)}(f_{2M-k})=\boldsymbol{Y}^{(p)}(f_k)$ 对 $k=0$，…，$M-1$ 成立，并且因为 $\boldsymbol{Y}^{(p)}(f_0)$ 和 $\boldsymbol{Y}^{(p)}(f_M)$ 当 M 增大的时候变得能够忽略它们的影响，我们能够证明，对于大的样本尺寸，$\boldsymbol{Y}^{(p)}$ 的离散小波变换和只有 $\boldsymbol{Y}^{(p)}(f_1)$，…，$\boldsymbol{Y}^{(p)}(f_{M-1})$ 的离散小波变换的统计性质对于公式化阈值水平是一样重要的(正如在下面见到的，我们能够忽略 $M-1$ 不是 2 的幂的复杂性). 有了这个简化途径，我们能够跟随 Moulin(1994) 来获得小波系数 $\{e_{j,t}\}$ 的统计性质. 因为由假设 $\{\varepsilon(f_k)\}$ 是独立同分布，并且离散小波变换是规范正交的，所以 $\{e_{j,t}\}$ 是不相关的，但是，由于 $\{\varepsilon(f_k)\}$ 与 $\log(\chi_2^2)$ 分布相关而不是高斯分布，$\{e_{j,t}\}$ 不是同分布的随机变量——它们的分布现在依赖于 j 层小波滤波器的效果.

借助于由 $e_{j,t}$ 而来的独立同分布的随机变量 $\{\varepsilon(f_k)\}$ 的线性组合项的数目，随 j 迅速增加(回忆 j 层的小波滤波器宽度由式(96a) $L_j\equiv(2^j-1)(L-1)+1$ 给出). 因此我们可求助于中心极限定理来证明 $e_{j,t}$ 的分布当 j 增长时候收敛到高斯分布. 既然 $\{e_{j,t}\}$ 是不相关的，Moulin(1994)指出，对于大的 j，$\{e_{j,t}\}$ 可以认为是独立随机变量. 在这种情况下，假设独立同分布高斯小波系数用小波阈值或收缩是合法的(我们在 10.5 节中讨论这种方法). 然而，在小而适当的尺度下，$e_{j,t}$ 的分布尾部不与高斯分布相近，但是当 j 增加时，则变得相当接近于高斯分布. 对这些尺度，Moulin(1994)用鞍点法来逼近 $e_{j,t}$ 分布的尾部特征，从而计算适当的阈值(对这种方法的背景，参见 Lugannani and Rice，1980；Daniels，1987；Reid，1988；和 Davison and Hinkley，1997).

与这种思想相符合，Moulin(1994)列出下面对阈值的逼近. 对每一个小波系数考虑零假设 H_0：$d_{j,t}=0$ 而不是相应的 H_1：$d_{j,t}\neq0$. 在零假设下，$W_{j,t}^{(p)}$ 与噪声项 $e_{j,t}$ 有相同的分布——由于它是用 $\log(\chi_2^2)$ 分布滤波一个随机变量得来——对

小而适当尺度将是反对称的. 我们不能在一个明显水平 α，当 $W_{j,t}^{(p)}$ 属于一个可接受区域$\left[\delta_{j,\frac{\alpha}{2}}^{(1)}, \delta_{j,\frac{\alpha}{2}}^{(u)}\right]$时，拒绝零假设，其中 $\delta_{j,\frac{\alpha}{2}}^{(1)}$ 和 $\delta_{j,\frac{\alpha}{2}}^{(u)}$ 分别是下阈值和上阈值使得

$$P\left[e_{j,t} < \delta_{j,\frac{\alpha}{2}}^{(1)}\right] = \frac{\alpha}{2} \quad 和 \quad P\left[e_{j,t} > \delta_{j,\frac{\alpha}{2}}^{(u)}\right] = \frac{\alpha}{2},$$

因此，当 H_0 是真的时，

$$P\left[\delta_{j,\frac{\alpha}{2}}^{(1)} \leqslant W_{j,t}^{(p)} \leqslant \delta_{j,\frac{\alpha}{2}}^{(u)}\right] = P\left[\delta_{j,\frac{\alpha}{2}}^{(1)} \leqslant e_{j,t} \leqslant \delta_{j,\frac{\alpha}{2}}^{(u)}\right] = 1 - \alpha. \tag{435}$$

由于对每一个 j 用了不同的小波滤波器，在 j 层上，上阈值和下阈值依赖于 $W_{j,t}^{(p)}$. 这个假设验证等价于对小波系数用下面的阈值法则：

$$W_{j,t}^{(t)} = \begin{cases} 0, & \delta_{j,\frac{\alpha}{2}}^{(1)} \leqslant W_{j,t}^{(p)} \leqslant \delta_{j,\frac{\alpha}{2}}^{(u)}; \\ W_{j,t}^{(p)}, & 其他. \end{cases}$$

α 的值可选择需要一个错误警告概率 P_F 取一个特别小的值，其中，由定义 P_F 至少是一个噪声小波系数落在接受区域外的概率. 因此

$$P_F = P\left[\bigcup_{j,t} (e_{j,t} > \delta_{j,\frac{\alpha}{2}}^{(u)} \quad 或 \quad e_{j,t} < \delta_{j,\frac{\alpha}{2}}^{(1)})\right] = 1 - P\left[\bigcap_{j,t} (\delta_{j,\frac{\alpha}{2}}^{(1)} \leqslant e_{j,t} \leqslant \delta_{j,\frac{\alpha}{2}}^{(u)})\right].$$

$\{e_{j,t}\}$是不相关的，但是如果我们通过将它们处理为独立系数进行逼近，上面就变成了

$$P_F = 1 - \prod_{j,t} P[\delta_{j,\frac{\alpha}{2}}^{(1)} \leqslant e_{j,t} \leqslant \delta_{j,\frac{\alpha}{2}}^{(u)}] = 1 - \prod_{j,t} (1-\alpha).$$

在简化的情况下，我们对 $Y^{(p)}(f_1)$，…，$Y^{(p)}(f_{M-1})$取接近于完全的离散小波变换(即在部分离散小波变换中有相当少的尺度系数)，J_0层变换大约有 M 个小波系数，逼近上面乘积给出项的数目. 因此我们就获得一个想要的错误警告概率(即 $P_F=0.1$)，置

$$\alpha = 1 - (1-P_F)^{1/M} \approx \frac{P_F}{M},$$

其中当 P_F 取很小的时候，使用逼近(基于 $\log(1-x) \approx -x$ 对很小的 x 成立)是很方便的.

• *例子*

让我们对 Gao(1993，1997)，Moulin(1994)用的三个高斯稳定过程使用周期图算法做例子.

[1]回忆，由定义，一个 p 阶自回归过程(AR(p))满足式(268e). Gao(1993，1997)考虑了 AR(24)过程，其系数$\{\phi_{24,n}\}$在表 272 中给出. 相应的谱密度函数由(268f)给出，而 $\sigma_\epsilon^2=1$，且 $\Delta t=1$，它表示为在图 438(还在图 273)最上面的细曲线. 我们能够用在 7.9 节中描述的方法形成这个过程的部分 X_0，…，X_{N-1} 的实现.

表 436

	j							
	1	2	3	4	5	6	7	8
$-\delta^{(l)}_{j,\frac{\alpha}{2}}$	7.825	7.031	6.228	5.750	5.460	5.287	5.182	5.118
$\delta^{(u)}_{j,\frac{\alpha}{2}}$	5.556	5.601	5.142	4.976	4.913	4.901	4.910	4.925

注：下阈值和上阈值 $\delta^{(l)}_{j,\frac{\alpha}{2}}$ 和 $\delta^{(u)}_{j,\frac{\alpha}{2}}$，$j=1, \cdots, 8$，对基于小波的对数周期图阈值使用 LA(8)离散小波变换．这里我们用 $\alpha=P_F/M$，$P_F=0.1$，$M=1\,024$ 的近似．为了方便，我们列表把 $-\delta^{(l)}_{j,\frac{\alpha}{2}}$ 替为 $\delta^{(l)}_{j,\frac{\alpha}{2}}$．注意随着 j 的增加，下阈值和上阈值在幅值上变得相近了，如希望的属于收敛到高斯性．

[2]Moulin(1994)定义一个 AR(2)过程，具体化为

$$X_t = \phi_{2,1} X_{t-1} + \phi_{2,2} X_{t-2} + \epsilon_t, \tag{436a}$$

其中 $\phi_{2,1} \equiv 0.97\sqrt{2}$ 和 $\phi_{2,2} \equiv -(0.97)^2$，$\{\epsilon_t\}$ 是具有零均值和单位方差的高斯白噪声．对于这个过程的谱密度函数能够通过式(268f)计算，而且在图 438 的中间的细曲线显示．再次，我们能够产生在 7.9 节中描述的实现方法．

[3]Moulin(1994)也定义了“典型的移动无线电通信”(MRC)谱密度函数，它是由两个带限，衰减的，移动无线电信号，一个背景白噪声，一个窄带有高斯形状干扰的谱密度函数的叠加．总的来说，谱密度函数是，对于 $0 \leqslant f \leqslant 1/2$，

$$S_X(f) = 10^{-3} + 0.2\mathrm{e}^{-(f-0.45)^2/(2.10^{-6})} + \left[1 - \left(\frac{f}{B_0}\right)^2\right]^{1/2} 1_{[0,B_0)}(f)$$

$$+ \left[1 - \left(\frac{f-f_0}{B_0}\right)^2\right]^{1/2} 1_{(f_0-B_0, f_0+B_0)}(f),$$

而有 $f_0=0.3$，$B_0=0.1$(如往常，对于 $f<0$，$S_X(f)=S_X(-f)$)．这个谱密度函数是在图 438 最底下的细曲线．这里，我们可以用在 7.8 节中描述的从已知谱的高斯过程中的方法来模拟．

对这些模型中的每一个，我们产生一千个模拟时间序列，长度 $N=2\,048$(因此 $M=1\,024$)并且，对于每个序列形成周期图．由 $P_F=0.1$，我们置 $\alpha=P_F/M \doteq 9.765\,6\times10^{-5}$，然后发现，通过鞍点法，下阈值和上阈值为 $\delta^{(l)}_{j,\frac{\alpha}{2}}$ 和 $\delta^{(u)}_{j,\frac{\alpha}{2}}$，$j=1, \cdots, 8$，用 LA(8)离散小波变换(阈值在表 436 中列出)．用这些与软阈值相关联的阈值，然后我们基于 $J_0=5, 6, 8$ 层部分 LA(8)离散小波变换计算 $\hat{S}_X^{(\mathrm{pt})}(\cdot)$(因此设置 J_0 的原因在于，分别用 64、32 和 8 的尺度系数用阈值后左边不受影响)．我们判断通过计算每一个谱密度函数估计的根均方差(RMSE)来判断它的质量，即

$$\left[\frac{1}{M+1}\sum_{k=0}^{M}\left[10\log_{10}(\hat{S}_X^{(\mathrm{pt})}(f_k)) - 10\log_{10}(S_X(f_k))\right]^2\right]^{1/2} \tag{436b}$$

(注意到这有分贝单位). 对于每一 J_0 层，我们在这一千个模拟序列计算平均根均方差(这些在图 446 下面用星号给出). 另外，对产生最小根均方差的水平我们已经绘出(图 438 中的粗曲线)谱密度函数估计一从这一千个这样的估计中一它们的根均方差最接近平均根均方差(对 AR(24)，AR(2)和移动无线电通信模型分别选择层为 $J_0=5$，8 和 5).

现在我们评论图 438 中的三个模型的谱估计.

[1]AR(24)过程的谱密度函数有一个动态的范围接近 90dB(由定义，这个范围是最大和最小的谱密度函数值的比例). 尽管 Gao(1997)在他的对数周期图的小波阈值中研究用了这个过程，实际上，这个周期图是有很坏的偏差，这是由于对样本尺寸 $N\leqslant 2\,048$ 和 $f\geqslant 0.4$ 的漏泄(参见图 273 或 Gao，1997 的图 1). 这种漏泄在谱密度函数估计的高频上已显出来，在图 438 最上面展示. 另外，对于 $f<0.4$，估计表示错了几个峰值，特别地，低估了主峰值大约 8dB 在接近于 $f=0.1$ 处. 由于这几点漏泄，其中的平均根均方差比其他两个模型大很多——参见图 446 的星型图形.

[2]图 438 中间的图表示对于 AR(2)过程的典型估计. 这个过程的谱密度函数比 AR(24)过程有更简单的结构，得到的谱密度函数估计总体上是相当好的(只有在主峰和高频有轻微的扭曲).

[3]最底下的图形表示对移动无线电通信过程的典型估计. 真正的谱密度函数比自回归过程谱密度函数有更灵敏的特性，但是它也有一个更小的波动范围. 估计没有 AR(2)模型光滑，但是更易于接受，但是在 $f=0.2$ 之前快速上升 30dB，而且有明显的扭曲.

• *谱估计方法评论*

用对数周期图作为对数谱密度函数估计的开始点有三个主要的缺点.

[1]首先，$\mathbf{Y}^{(p)}$ 中"噪声"是 $\log(\chi_2^2)$ 分布的，所以有明显不同于高斯分布的尾巴，因此我们不能对所有层的小波系数用简单的基于高斯的阈值. Moulin(1994)用的阈值水平是样本尺寸，层(尺度)和小波依赖——没有可提供的软件——需要一些努力来计算.

[2]其次，不考虑小波阈值，周期图可给出非常差的谱密度函数估计，无论何时真的谱密度函数有一个大的波动范围或是快速变化(甚至样本尺寸可能认为是大的). 这种缺乏性产生于周期图不能对边漏泄产生很好的防范效果. 这点在 Percival and Walden(1993，6.3 节)中详细讨论.

[3]为了保持周期化不相关的性质，我们必须在其傅里叶频率处取样，这要求在时间序列中观察数目的必须是 2 的幂. 如果这种假设没有满足，一个简单的——但不提倡的——改变是取一个网格频率粗于傅里叶频率的周期图，这对于

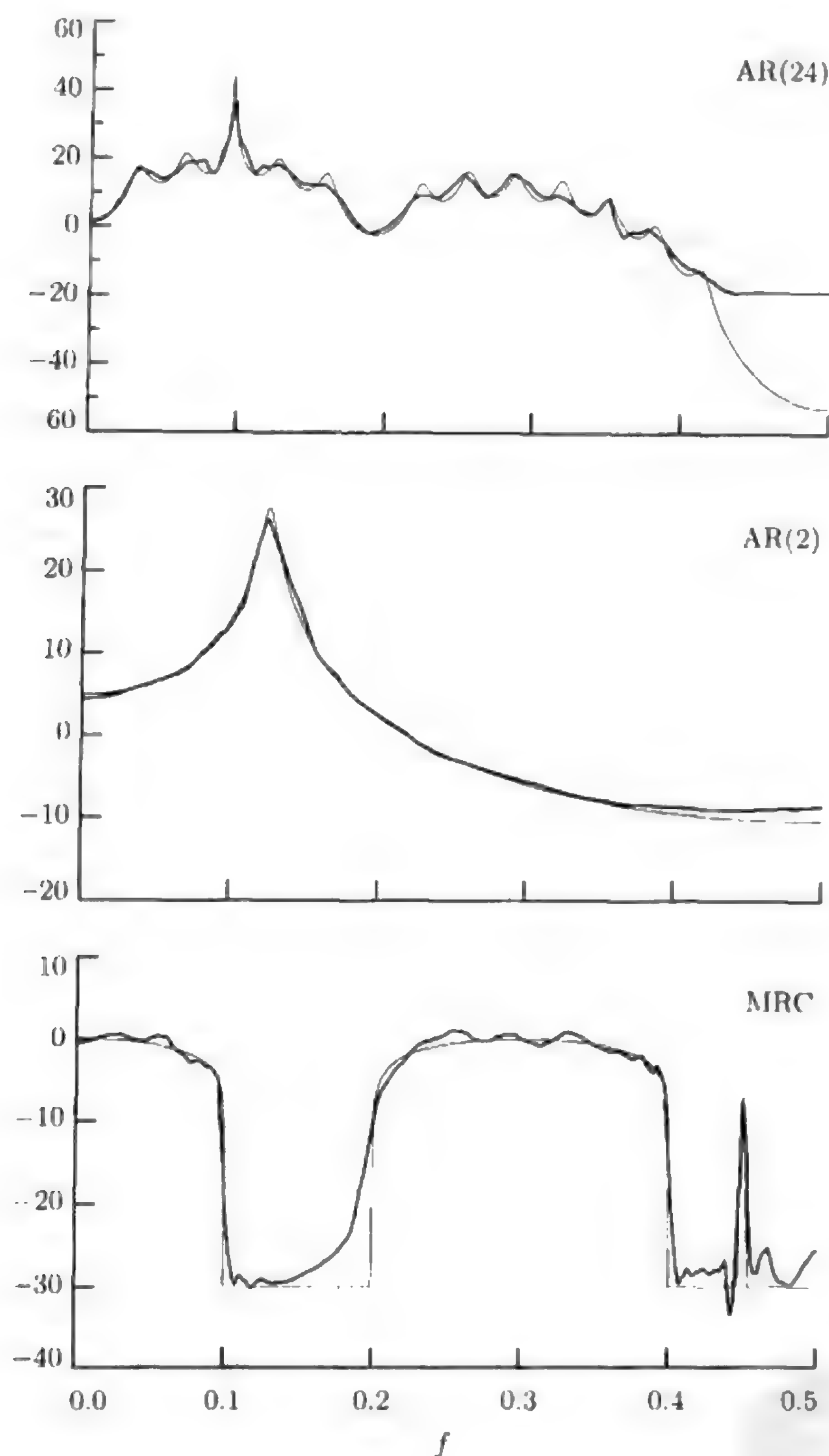


图 438 对 AR(24)，AR(2)和移动无线电通信过程基于周期图的估计谱密度函数(粗曲线)和真正的谱密度函数(细曲线)(详情参见正文)

不计某些时间序列的值是很有效的(这种变化能够变得更有吸引力，用一个"折叠部分平均"方法——例如，见 Percival and Walden，1993，6.17 节).

在 10.7 节中我们给出了基于多锥的谱密度函数估计围绕这三个实际缺点的另一个算法.

10.6 节的评论与扩展

[1]Gao(1993，1997)发展了对阈值对数周期图算法的一个稍微复杂的逼近．他产生了一个简单的阈值水平 $\delta_{j,M}$使得当 $M\to\infty$ 时(即 $N\to\infty$ 时)，

$$P_F = \boldsymbol{P}\Big[\bigcup_{j,t}(|e_{j,t}|>\delta_{j,M})\Big]\to 0.$$

另一个形式可写为

$$P_F = \boldsymbol{P}\Big[\bigcup_{j}(\max_t|e_{j,t}|>\delta_{j,M})\Big]\to 0.$$

因此至少一个噪声小波系数超过阈值 $\delta_{j,M}$的概率，随样本尺寸的增加而趋于零．阈值定义为

$$\delta_j^{(G)} = \max\{\alpha_j\log(M),\pi\sqrt{[\log(M)/3]}\},$$

其中系数$\{\alpha_j\}$由 Gao(1993，1997)对 $j=1$，…，10 列表为

$$\{\alpha_1,\cdots,\alpha_{10}\} = \{1.29,1.09,0.92,0.77,0.65,0.54,0.46,0.39,0.32,0.27\}.$$

这对于普通用的紧支撑的规范正交小波设计是有效的．现在 $\delta_j^{(G)}=\alpha_j\log(M)$，当 $\alpha_j\geqslant\pi/\sqrt{[3\log(M)]}$时．例如，如果 $M=512$，这个条件变为 $\alpha_j\geqslant 0.73$．对于这个样本尺寸的阈值在四个更小的尺度处取 $\delta_j^{(G)}=\alpha_j\log(M)$的形式，然后在粗尺度 $j\geqslant 5$ 处为 $\delta_j^{(G)}=\pi\sqrt{\log(M)/3}$．这对于阈值无疑是个简单的逼近，但无论如何仍用对数周期图作为基本的谱密度函数估，因为它实质上是不能抑制边缘漏泄的．

[2]在 10.2 节，我们注意到渐进的论证是用来证明，例如，在我们收集更多的数据时，统一阈值要求取样在确定的区间上的信号更精细些．对于，这里考虑的(还有在下部分)谱估计问题取样算法实际上是非常自然的．我们把对数谱作为一个信号，因此能观察确定频率区间$[0, f_N]$．随着时间序列的样本尺寸 N 的增加，我们可通过在网格频率跨度为 $\Delta f\equiv 1/(N\Delta t)$上的对数周期图取样我们的信号．对数周期图取样区间 Δf 当 $N\to\infty$的降低到零．所以我们对信号的取样变得越来越细，如对于渐近说明中需要的．

[3]为了公式化本节的理论，我们假设 $E\{X_t\}=0$．在实际使用中，我们不知道$\{X_t\}$的一个先验均值．通常的步骤是，使用 X_0，…，X_{N-1}的取样均值 $\overline{X}$ 来估计未知的过程均值．之后，我们能够用 $X_t-\overline{X}$ 代替 X_t来计算周期图，给予 $\hat{S}_X^{(p)}(f_0)=0$，因此 $Y^{(p)}(f_0)=-\infty$．为了克服这些困难，我们重新定义式(433b)的 $\boldsymbol{Y}^{(p)}$，使它变为

$$Y^{(p)}(f_1),Y^{(p)}(f_2),\cdots,Y^{(p)}(f_M),Y^{(p)}(f_M),\cdots,Y^{(p)}(f_2),Y^{(p)}(f_1),$$

为了保持在式(140)中的反射边界条件．注意因为 $Y^{(p)}(f_{2M-k})=Y^{(p)}(f_k)$对于 $k=1$，…，$M-1$ 成立，上面等价于从 $Y^{(p)}$中除去 $Y^{(p)}(f_0)$，然后复制 $Y^{(p)}(f_M)$来获得直到 2 的幂的向量大小．

10.7 相关高斯小波系数

在 10.6 节我们介绍了通过周期图小波阈值的谱估计. 由于噪声小波系数 $\{e_{j,t}\}$是不相关的, 所以不是同分布的, 导致了更复杂的阈值法则. 本节, 我们研究噪声小波系数$\{n_{j,t}\}$是相关的, 却逼近于高斯分布的情形. 这里感兴趣的应用涉及谱估计, 但这次通过多锥小波阈值谱估计而利用 K 个正弦锥 $\{a_{n,t},\ t=0,\ \cdots,\ N-1\}$, $n=0,\ \cdots,\ K-1$(参见 7.5 节式(274a)).

• **谱估计算法概要**

如在 10.6 节中, 令 X_0, …, X_{N-1}是一个零均值的平稳过程的一部分, 有谱密度函数 $S_X(\cdot)$和取样间隔 Δt. 令 $2M=2^{J+1}$是一个大于或等于样本尺寸 N 的 2 的幂. 我们通过"填塞零"将锥的时间序列$\{a_{n,t}X_t,\ t=0,\ \cdots,\ N-1\}$扩展到长度为 $2M$, 产生 K 个序列

$$\{a_{n,0}X_0,\cdots,a_{n,N-1}X_{N-1},\underbrace{0,\cdots,0}_{2M-N\text{个}0}\},\quad n=0,\cdots,K-1.$$

对这些序列应用标准快速傅里叶变换算法, 在 $M=2^J$ 频率 $f_k\equiv k/(2M\Delta t)$, $k=1$, …, $M-1$ 能够容易计算对数多锥谱密度函数坐标:

$$Y^{(\mathrm{mt})}(f_k)=\log(\hat{S}_X^{(\mathrm{mt})}(f_k))-\psi(K)+\log(K)=\log(S_X(f_k))+\eta(f_k),$$

其中 $\psi(\cdot)$是一个 Γ_3函数. 上面说到, 对数多锥估计(加已知常数)能够写为一个信号(真对数谱)加噪声, 其中一假设我们用适当数目的锥($K\geqslant5$)一噪声是近似高斯的, 而有零均值和已知方差 $\sigma_\eta^2=\psi'(K)$(参见 7.5 节, 如以前的, $\psi'(\cdot)$是 Γ_3 函数). 由于没有必要估计在傅里叶频率 $k/(N\Delta t)$的谱, 观察的数目就不必是 2 的幂. 注意其中用零填塞是很有理由的, 因为我们假设我们的过程$\{X_t\}$有零均值.

我们能够用向量记号表述这些结果为

$$\begin{bmatrix}Y^{(\mathrm{mt})}(f_0)\\ \vdots\\ Y^{(\mathrm{mt})}(f_{M-1})\end{bmatrix}=\begin{bmatrix}\log(S_X(f_0))\\ \vdots\\ \log(S_X(f_{M-1}))\end{bmatrix}+\begin{bmatrix}\eta(f_0)\\ \vdots\\ \eta(f_{M-1})\end{bmatrix}\equiv\boldsymbol{D}+\boldsymbol{\eta}\qquad(440)$$

(参见式(433a)). 如对于对数周期图, 这可熟练地扩展到式(276)的模型, 包括 $f=f_0=0$ 这种情况, 虽然这不严格正确; 再次, 这一个附加项的影响对大的 N 可忽略. 对上述应用完全离散小波变换给出

$$W_{j,t}^{(\mathrm{mt})}=d_{j,t}+n_{j,t}\quad\text{和}\quad V_{J,0}^{(\mathrm{mt})}=d_{J+1,0}+n_{J+1,0}$$

对 $j=1$, …, J 和 $t=0$, …, M_j-1 成立, 其中 $V_{J,0}^{(\mathrm{mt})}$是信号和噪声分量 $d_{J+1,0}$和 $n_{J+1,0}$的单个尺度系数, 且 $M_j\equiv M/2^j$.

• 系数和阈值统计性质的说明

我们把 η 作为一个多元的高斯向量模型，具有零均值，协方差矩阵为

$$\Sigma_\eta \equiv \begin{bmatrix} s_\eta(f_0) & \cdots & s_\eta(f_{\frac{M}{2}-1}) & s_\eta(f_{\frac{M}{2}}) & s_\eta(f_{\frac{M}{2}-1}) & \cdots & s_\eta(f_1) \\ s_\eta(f_1) & \cdots & s_\eta(f_{\frac{M}{2}-2}) & s_\eta(f_{\frac{M}{2}-1}) & s_\eta(f_{\frac{M}{2}}) & \cdots & s_\eta(f_2) \\ s_\eta(f_2) & \cdots & s_\eta(f_{\frac{M}{2}-3}) & s_\eta(f_{\frac{M}{2}-2}) & s_\eta(f_{\frac{M}{2}-1}) & \cdots & s_\eta(f_3) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ s_\eta(f_1) & \cdots & s_\eta(f_{\frac{M}{2}}) & s_\eta(f_{\frac{M}{2}-1}) & s_\eta(f_{\frac{M}{2}-2}) & \cdots & s_\eta(f_0) \end{bmatrix}, \quad (441a)$$

而 $s_\eta(f_k)$是在式(277)中定义. 这个循环的协方差矩阵是对称半正定的. 它的分量与在 7.5 节的讨论的 $\eta(f_k)$一致，$\eta(f_k)$使 f_k从 0 与 f_N充分远；对于其他频率(在数目是相当小的，随 N 的增加重要性降低)，它提供了一种下面很好运用的方便的数学结构. 矩阵 Σ_η能够被认为是对 η 协方差结构的一种更自然的假设的逼近，即对称的 Toeplitz 协方差矩阵 Σ'_η. 其第一行由 $s_\eta(f_k)(k=0, \cdots, M-1)$给出(一个 Toeplitz 矩阵定义为，其分量 $\Sigma'_{\eta,l,m}=s_\eta(f_{l-m})$即沿着一个对角是一个单个值，因此，第一行决定了整个矩阵). 因为式(277)，矩阵 Σ_η几乎是一个带状的 Toeplitz 矩阵：它有 $2k+1$ 个非零对角线，还有 $K(K+1)/2$ 个非零“带外”分量在右上角和左下角. 这些相对少的“带外”元素用来区分 Σ_η 和 Σ'_η.

因为在式(441a)中的协方差矩阵 Σ_η是循环的，我们可引起在时间序列分析中的著名结果(参见 Fuller, 1996, p. 151)写 $\Sigma_\eta=\mathcal{F}^{H}D\mathcal{F}$，其中 $\mathcal{F}$ 是一个 $M\times M$ 阶正交离散傅里叶变换矩阵(参见 3.4 节)；“H”代表埃尔米特转置；D 是对角元为$\{S_k: k=0, \cdots, M-1\}$的对角矩阵，它是 Σ_η的第一行的离散傅里叶变换. 下面的结果给我们一个容易计算方差 $n_{j,t}$的简单方法.

练习[441] 对 $M=2^J$，令 $\mathbf{n}=\mathcal{W}\eta$ 是高斯平稳随机过程的离散小波变换系数，其协方差矩阵为在式(441a)中给出的 $M\times M$ 阶 Σ_η，并且令$\{n_{j,t}: t=0, \cdots, M_j-1\}$是与 j 层有关的 $\mathbf{n}$ 个元素($M_j\equiv M/2^j$). 采取基于宽度为 $L\leqslant M$ 的小波滤波器离散小波变换，证明

$$\operatorname{var}\{n_{j,t}\}=\frac{1}{M}\sum_{k=0}^{M-1}S_k\,\mathcal{H}_j\left(\frac{k}{M}\right)\equiv\sigma_j^2, \quad (441b)$$

其中，照例 $\mathcal{H}_j(\cdot)$是第 j 层小波滤波器$\{h_{j,l}\}$的平方增益函数(特别注意，$\operatorname{var}\{n_{j,t}\}$依赖于 j 而不依赖于 t). ◁

式(441b)能够容易地计算，对层依赖的方差 σ_j^2给出我们知道的值.

这个方差是怎样随层 j 的变化而变化的? 对于 $M=2^J=N/2$ 而且在温和条件 $K<M/2$ 下的特别的哈尔离散小波变换，可表示为

$$\sigma_1^2<\sigma_2^2<\cdots<\sigma_J^2<\bar{\sigma}_{J+1}^2, \quad (441c)$$

其中 $\bar{\sigma}_{J+1}^2$ 是噪声尺度系数 $n_{J-1,0}$ 的方差(详情参见 Walden 等，1998). 式(441b)能够用于数值上说明上面的结论对基于其他的 Daubechies 小波的离散小波变换仍然适用(参见练习[10.9]).

关于噪声小波系数方差的最后问题的回答是，它们与 $\sigma_\eta^2=\psi'(K)$ 怎样相关. 由于 $\mathcal{W}$ 是一个正交变换，$\mathrm{tr}\{\Sigma_\eta\}=\mathrm{tr}\{\mathcal{W}\,\Sigma_\eta\,\mathcal{W}^{\mathrm{T}}\}=\mathrm{tr}\{\Sigma_\eta\}$，其中 $\mathrm{tr}\{\cdot\}$ 表示矩阵的迹(即对角分量的和). Σ_η 的对角由共同的元素 $s_\eta(f_0)=\sigma_\eta^2$ 组成，使 $\mathrm{tr}\{\Sigma_\eta\}=M\sigma_\eta^2=2^J\sigma_\eta^2$. 从式(441b)中我们知道 $\mathcal{W}\,\Sigma_\eta\,\mathcal{W}^{\mathrm{T}}$ 的对角具有形式

$$\underbrace{\sigma_1^2,\cdots,\sigma_1^2}_{\frac{M}{2}\uparrow},\underbrace{\sigma_2^2,\cdots,\sigma_2^2}_{\frac{M}{4}\uparrow},\cdots,\sigma_J^2,\bar{\sigma}_{J+1}^2.$$

因此

$$\sigma_\eta^2=\frac{\bar{\sigma}_{J+1}^2}{2^J}+\sum_{j=1}^{J}\frac{\sigma_j^2}{2^j},$$

与式(441c)结合，表明

$$\sigma_1^2<\sigma_\eta^2<\bar{\sigma}_{J+1}^2. \tag{442}$$

计算表明，在哈尔离散小波变换和 $M=N/2$ 这种特别的情况下，实际上当 $K=5,\cdots,9$ 时，$\sigma_2^2<\sigma_\eta^2<\sigma_3^2$；当 $K=10$ 时，$\sigma_\eta^2\approx\sigma_3^2$；当 $K=11,\cdots,41$(这包含了 K 在实用中所有可能的锥的数目)时，$\sigma_3^2<\sigma_\eta^2<\sigma_4^2$.

为了确定 σ_j^2 随着 j 的增加而增加，更重要的是，为了证实由式(277)，(440)和(441a)定义的随机模型是正确的，我们采取了一个模拟学习，在模拟学习中下面的步骤重复了一千次.

[1]从具体的高斯平稳过程中产生取样 $X_0,\cdots,X_{N-1}$，其中 $N=2\,048$，$M=N/2$.

[2]使用 $K=10$ 的正弦曲线锥计算多锥估计.

[3]基于长度 $L=2$，4，8 和 16 的小波滤波器式(440)中的 η 的离散小波变换计算，产生小波系数 $\{n_{j,t}\}$.

[4]最后，在 $j=1$，2，3 和 4 层，小波系数 $\{n_{j,t}\}$ 的标准离差由 $\hat{\sigma}_j^2\equiv\frac{1}{M_j}\sum_{t=0}^{M_j-1}n_{j,t}^2$ 的平方根来估计.

使用了两个不同的过程，白噪声过程和 AR(2)过程在式(436a)中定义，但提供的结果在本质上是相同的，因此这里我们只通报了 AR(2)的情况. 图 443 中的盒图绘出了每个小波滤波器每一层一千个标准离差. 盒图中是这样定义的：细水平线在数据的中位数通过盒子，盒子的上底和下底是在上盒下的四分位数，由盒子延伸出来的垂直线是为了指出在数据的标准离差内盒的垂直线延长超过了每个盒子，定义为里面范围的 1.5 倍. 估计 $\hat{\sigma}_j$ 的方差随层数 j 的增加而增加. 延伸到每个盒子外面的粗的垂直线表明 σ_j 的值是从式(441b)中产生的. “标称的”的标准离差 $\sigma_\eta\doteq\sqrt{\psi'(K)}=\sqrt{\psi'(10)}\doteq0.32$ 被记为旋转于每个图形宽度的细线. 当我们用

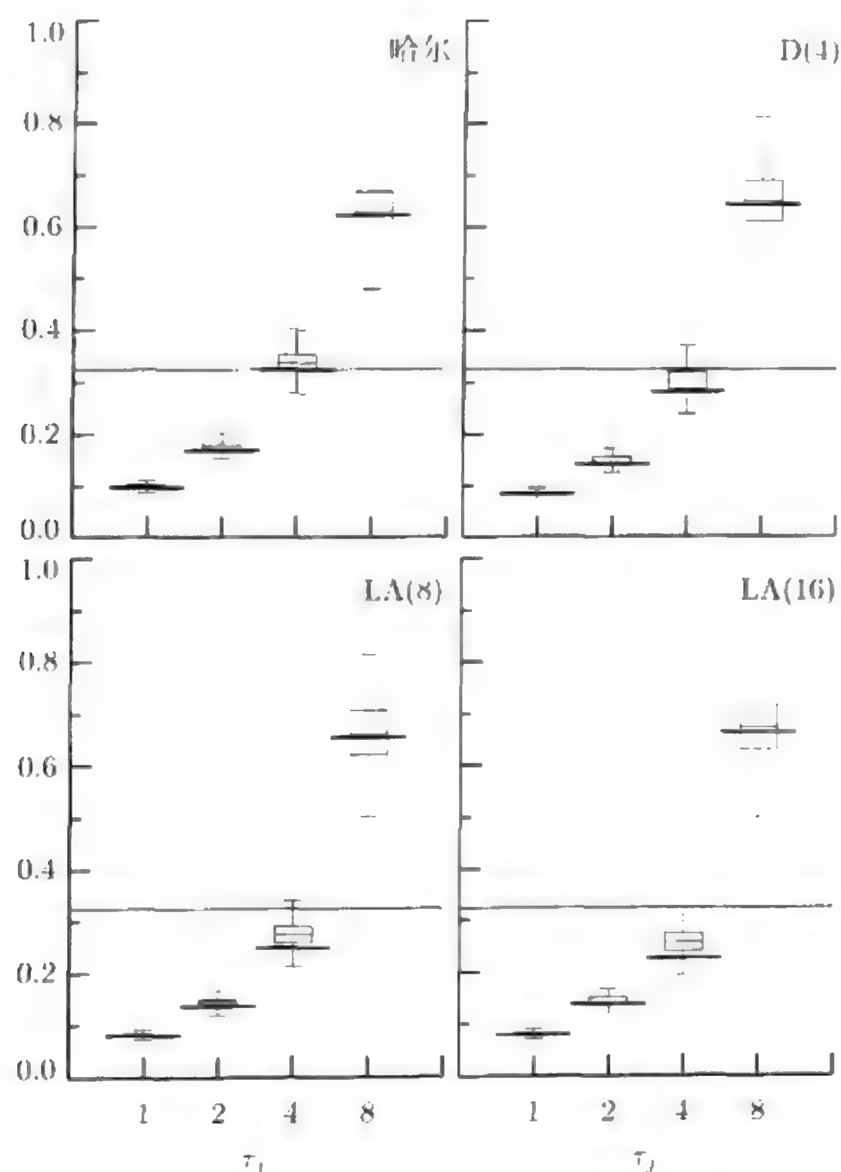


图 443　用不同的小波滤波器对 $N=2\,048$ 和 $K=10$ 从 AR(2)过程中产生的小波系数 $n_{j,t}$，在 $j=1$，2，3，4 层的估计标准离差 $\hat{\sigma}_j$ 的盒图. 从盒中延伸出来的水平实线表明从式(441b)中产生的值 σ_j. 在每个图中"标称"的标准离差 $\sigma_\eta=\sqrt{\psi'(10)}\doteq 0.32$ 被标识为穿过盒子的水平实线

D(4)、LA(8)或者 LA(16)小波滤波器时，这对 1、2、3 层超过了取样 $\hat{\sigma}_j$ 的均值. 对哈尔小波滤波器，标称的标准离差对 1、2 层差超过了 $\hat{\sigma}_j'$ 的取样均值. 因此，图 443 解释了式(441c)和导致对式(277)、(440)和(441a)逼近的信任.

现在，让我们重新将注意力转移到阈值小波系数 $\{W_{j,t}^{(mt)}\}$. 对所有变换层 j，假设我们选择阈值层为确定的，相应于统一阈值

$$\delta^{(u)}\equiv\sqrt{[2\sigma_\eta^2\log(M)]}. \tag{443}$$

从图 443 和式(442)中易见，相关性相应于小波阈值示例来抑制小尺度(低于 j)

“噪声尖峰”而保持信号的粗尺度系数(高于 j)不减弱(注意图 443 中的例子，我们有 $\sigma_\eta \geqslant \sigma_j$　$j=1$，2，3 和 $\sigma_4 > \sigma_\eta$). 我们称这个方法为*层独立(统一)*阈值. 注意 $\sigma_\eta = \sqrt{\psi'(K)}$ 是已知的，仅依赖于我们选择的锥数 K 的数目.

一个替代的方法是让阈值随变换层而改变，而对每一层仍用统一阈值公式：

$$\delta_j^{(u)} \equiv \sqrt{[2\sigma_j^2 \log(M)]}. \tag{444a}$$

这就是*层—依赖(统一)*阈值. Johnstone and Silverman(1997)给出了这样的阈值的例子，但是他们不得不由观察小波系数估计每一个变换层的噪声小波系数的标准离差(参见本节的评论与扩展). 我们已经处于有力的位置，从式(441b)中确定 σ_j^2.

最后，为了得到一个实际的过程，考虑到离散小波变换内在的循环假设，我们如上节讨论的，取式(440)展开形式的离散小波变换，即

$$\boldsymbol{Y}^{(mt)} \equiv \begin{bmatrix} Y^{(mt)}(f_0) \\ \vdots \\ Y^{(mt)}(f_{2M-1}) \end{bmatrix} = \begin{bmatrix} \log(S_X(f_0)) \\ \vdots \\ \log(S_X(f_{2M-1})) \end{bmatrix} + \begin{bmatrix} \eta(f_0) \\ \vdots \\ \eta(f_{2M-1}) \end{bmatrix} \equiv \boldsymbol{D} + \boldsymbol{\eta} \tag{444b}$$

(参见式(433b)).

• *全谱估计算法*

现在我们可以十分详细地说明谱估计算法，即

[1]X_0，…，X_{N-1} 时间序列为一个零均值平稳过程的一部分的实现；

[2]K，锥的数目；

[3]M，2 的幂满足 $2M \equiv 2^{J+1} \geqslant N$；

[4]J_0，满足 $J_0 \leqslant J$ 的层；

[5]选择阈值方法(例如硬、中、软阈值).

我们用下面的方式构造谱估计 $\hat{S}_X^{(mthr)}$.

[1]我们从计算多锥估计的对数算法开始. 对每一个正弦锥 $\{a_{n,t}\}n=0$，…，$K-1$，如在式(274a)中定义的，我们形成锥序列 $a_{n,0}X_0$，…，$a_{n,N-1}X_{N-1}$，附加 $2M-N$ 个零到这个序列，并且用一个 2 的幂的标准快速傅里叶算法形成特征谱

$$\hat{S}_{X,n}^{(mt)}(f_k) = \Delta t \left| \sum_{t=0}^{2M-1} a_{n,t} X_t e^{-i2\pi f_k t \Delta t} \right|^2, \ k = 0, \cdots, 2M-1,$$

其中 $f_k = k/(2M\Delta t)$ 而且 $a_{n,t}X_t \equiv 0$，$t \geqslant N$. 这 K 个特征谱平均形成如 7.5 节中的多锥估计 $\hat{S}_X^{(mt)}(f_k)$. 然后，我们形成 $Y^{(mt)}(f_k) \equiv \log(\hat{S}_X^{(mt)}(f_k)) - \psi(K) + \log(K)$

并且将这些 $2M$ 个值放在列向量 $Y^{(\mathrm{mt})}$ 中，如每个式(444b).

[2]下面，我们使用部分离散小波变换到 J_0 层. 我们获得部分离散小波变换系数 $\boldsymbol{W}_1^{(\mathrm{mt})}$，…，$\boldsymbol{W}_{J_0}^{(\mathrm{mt})}$ 和 $\boldsymbol{V}_{J_0}^{(\mathrm{mt})}$. 逐个小波系数分量能够写成

$$W_{j,t}^{(\mathrm{mt})}=d_{j,t}+n_{j,t},\quad j=1,\cdots,J_0,\quad t=0,\cdots,M_{j-1}-1.$$

在 $\boldsymbol{V}_{J_0}^{(\mathrm{mt})}$ 中的 M_{J_0} 个尺度系数完全没有管.

[3]我们把 $\boldsymbol{W}_1^{(\mathrm{mt})}$，…，$\boldsymbol{W}_{J_0}^{(\mathrm{mt})}$ 取阈值得到 $\boldsymbol{W}_1^{(\mathrm{t})}$，…，$\boldsymbol{W}_{J_0}^{(\mathrm{t})}$，经过

(a)对所有的 $j=1$，…，J_0 层应用由式(443)的层独立的阈值得到小波系数，或者

(b)分别对 $j=1$，…，J_0 层用式(444a)层依赖阈值.

[4]我们用逆变换 $\boldsymbol{W}_1^{(\mathrm{t})}$，…，$\boldsymbol{W}_{J_0}^{(\mathrm{t})}$ 和 $\boldsymbol{V}_{J_0}^{(\mathrm{mt})}$ 得到式(444b)的经由 $\hat{\boldsymbol{D}}^{(\mathrm{t})}$ 的 $\boldsymbol{D}$ 的估计式. 令 $\hat{D}_k^{(\mathrm{t})}$ 表示 $\hat{\boldsymbol{D}}^{(\mathrm{t})}$ 的第 k 个分量，注意 $\exp(\hat{D}_k^{(\mathrm{t})})$ 给出 $S_X(f_k)$ 的一个估计. 现在我们可以基于下面的阈值的多锥形成谱估计：

$$\hat{S}_X^{(\mathrm{mthr})}(f_k)\equiv\begin{cases}\mathrm{e}^{\hat{D}_0^{(\mathrm{t})}}, & k=0;\\ \dfrac{1}{2}\left(\mathrm{e}^{\hat{D}_k^{(\mathrm{t})}}+\mathrm{e}^{\hat{D}_{2M-k}^{(\mathrm{t})}}\right), & k=1,\cdots,M-1;\\ \mathrm{e}^{\hat{D}_M^{(\mathrm{t})}}, & k=M.\end{cases}$$

如果 $\{X_t\}$ 不能假定有零均值(实际应用中通常的情况). 我们在步骤[1]中用 $X_t-\overline{X}$ 替换 X_t，其中 $\overline{X}$ 为样本均值(如上节的评论与扩展的第[3]条中，以这个方法重新定 X_t 的中心使得周期图在零频率等于零的效果. 相反，一个多锥估计通常不易影响，因此我们可不调整 $\boldsymbol{Y}^{(\mathrm{mt})}$).

• *例子*

让我们对同样的 AR(24)、AR(2)和移动无线电通信模型应用层独立和层依赖阈值算法，以用于研究基于周期图的算法. 对于这些模型中的每一个，我们都产生长度为 $N=2\,048$ 的一千个模拟时间序列，然后形成基于 $K=10$ 的正弦锥上的多锥的谱密度函数估计. 这些估计的带宽是 $(K+1)/(N+1)\approx0.005\,4$，这与三个模型中峰值的宽度有极小的相关性. 之后，我们基于 LA(8)小波滤波器计算 $\hat{S}_X^{(\mathrm{mthr})}(\cdot)$，而结合层独立和层依赖的硬、中、软阈值，以及 $J_0=5$，6，8. 我们像在式(436b)中用计算根均方差判断每一个谱密度函数估计的品质，但是用 $\hat{S}_X^{(\mathrm{mthr})}(f_k)$ 代替 $\hat{S}_X^{(\mathrm{pt})}(f_k)$. 对每一个类型阈值的这些根均方差的平均和在一千个模拟中 J_0 的选择，在图 446 中用的是不同类型的连线(星型线表示对于周期图方法的相似的根均方差). 另外，根均方差接近于平均的根均方差的谱密度函数估计在图 447、448 和 449 中绘出. 作为模型的确定性选择的一个代表性的谱估计，阈值类型和 J_0.

现在让我们评论一下对于三种模型中的每一个谱估计的性质.

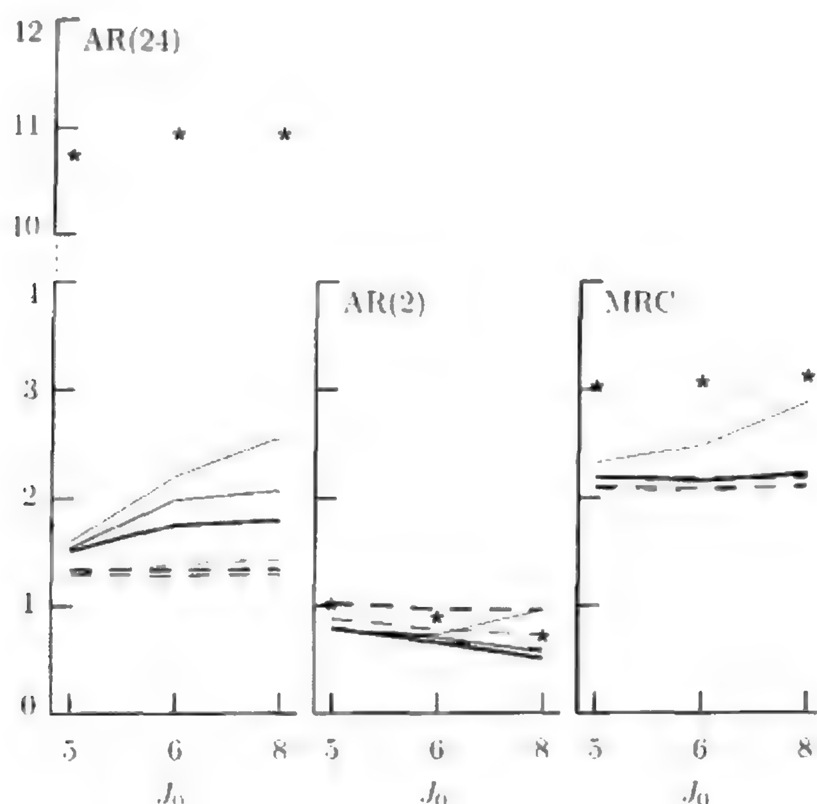


图 446 对于 AR(24)(左边图形)、AR(2)(中间)和移动无线电通信模型(右边)的根均方差一千个模拟的平均值. 在每一个图形中，对于层依赖的多锥方法的硬阈值(实粗曲线)、中阈值(实中等线)和软阈值(实细线)以及对于层独立的硬阈值(虚粗曲线)、中阈值(虚中曲线)和软阈值(虚细线)的平均的根均方差(以分贝算)也都绘出来了. 考虑了 J_0 层的三个值即 5，6 和 8，相应于阈值后不变的 64，32 和 8 尺度系数. 星形表示了对于基于周期图方法的平均根均方差. 对所有的情况，序列的长度是 $N=2\,048$，并且我们用的是 LA(8)小波计算离散小波变换

[1]与周期图不同，多锥算法抑制漏泄，并且当与小波阈值结合的时候产生了没有漏泄的光滑对数谱密度函数估计，如图 447 所示. 这个图表示了对层独立软阈值(下面的图形)和层依赖硬阈值(上面的图形)估计的代表方法，其中 $J_0=5$(这些选择给出了对层独立和层依赖的阈值参数结合的最小平均根均方差；参见图 446 左边的图形). 两个估计都是好的，易于获得比相应的基于周期图的估计有更好的峰值(图 438 上面)；然而，层独立估计可能被判断为优于层依赖估计谱密度函数的三个最高频率峰值的估计. 层依赖估计证明了小尺度的噪声系数在阈值下没有变得稀薄.

[2]图 448 表现了对层独立软阈值和层依赖硬阈值的 AR(2)过程的代表性估计，两个都有 $J_0=8$ 层(这些选择又给出了对层独立和层依赖阈值参数结合的最小平均根均方差；参见图 446 中间的图形). 两个估计是相当好的. 层依赖方法

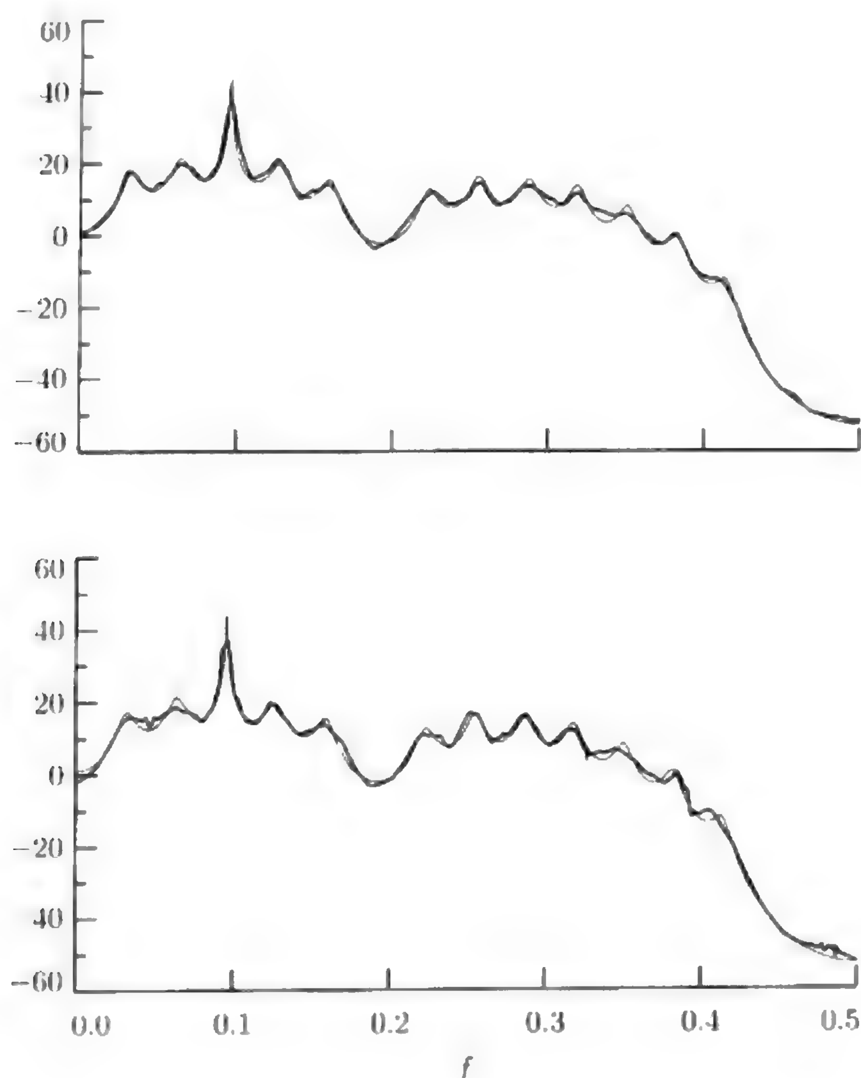


图 447　对 AR(24)过程估计的谱密度函数(粗曲线)和真正的谱密度函数(细曲线). 谱密度函数估计有代表性的是因为他们有接近在一千个模拟中的平均根均方差. 上面的图形是层独立的软阈值，而且下面的图形是层依赖的硬阈值，在两种情况下 $J_0=5$. 模拟序列的长度为 $N=2\,048$，我们用的是 LA(8)离散小波变换

(下面图)再次证明了小尺度的噪声系数由阈值不减弱，但是对估计谱密度函数的峰值的高度起很少的作用. 即使这个过程周期图本质上对适当尺度的也漏泄. 实际上，基于多锥估计比那些基于周期图的算法有某些更小平均的根均方差.

[3]图 449 表示对移动无线电通信模型层独立软阈值和层依赖硬阈值的代表性估计，$J_0=6$(接近前边基于图 446 右边图的研究). 不但在谱密度函数的带限部分升了和降了 30dB，而且在干扰 $f=0.45$ 的窄带用两种方法都可以很好地估计，而层独立估计(上面图)在获取峰值高度时更好点(如图 438 所示，基于周期图的估计也不能获得在谱密度函数中最尖的迁移). 层依赖估

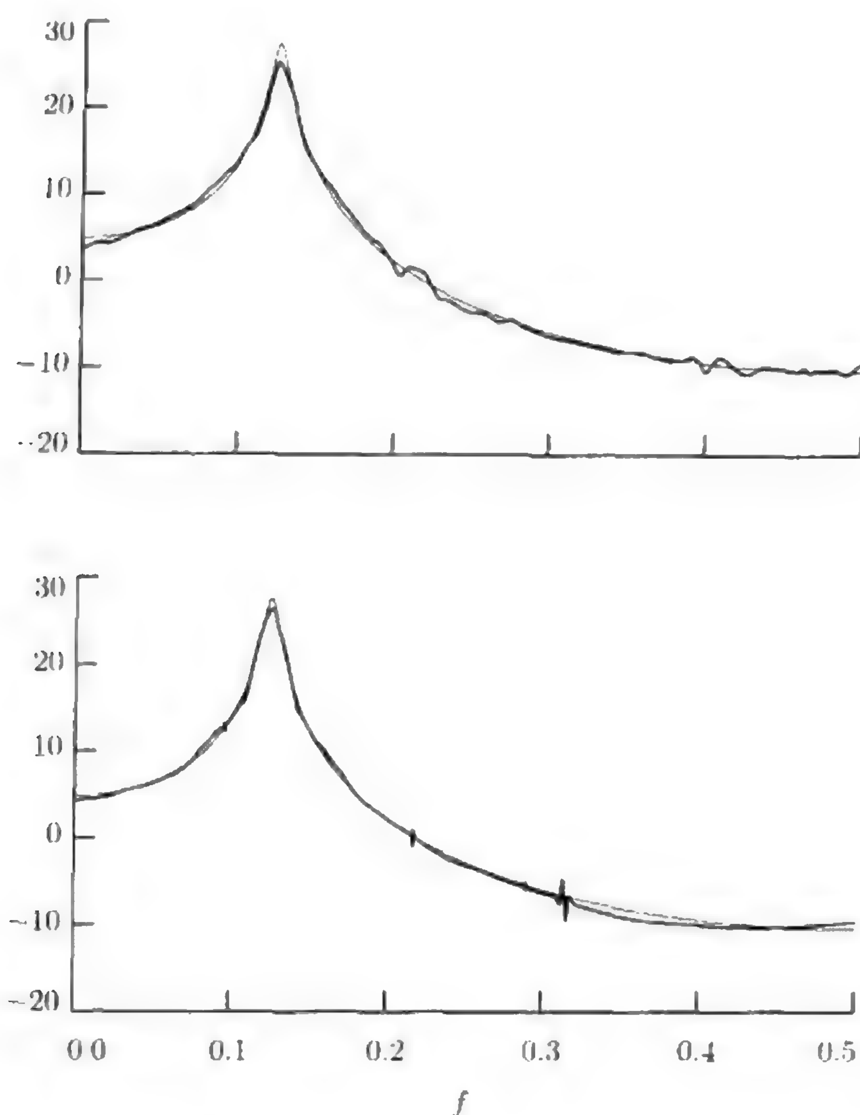


图 448 对 AR(2)过程估计的谱密度函数(粗曲线)和真正的谱密度函数(细曲线). 设计和参数参见图 447

计(下面图)也证明了阈值下的小尺度噪声系数没有减小. 基于多锥估计相比于基于周期图估计在根均方差上有明显的减少(参见图 446 右边图形).

注意最好的 J_0 层对三个模型是不同的, 而谱估计的根均方差随 J_0 的选择而有点变化. 如图 446 所示; 然而, 假设软阈值用于层独立统一阈值, 硬阈值用于层依赖统一阈值, J_0层的根均方差的变化不能使在视觉表现上的相应的谱估计有明显的不同.

最后, 让我们对层依赖和层独立阈值进行议论. 图 446 实曲线表示前者的表现, 而虚线则表示后者. 层独立阈值对所有的 AR(24)和移动无线电通信模型很好. 对 AR(2)用软阈值是可以同能提供的最好的层依赖阈值相比的. 对于为何层独立阈值能表现这么好的深入看法可从图 443 中看出一些. 在图形中缠绕平

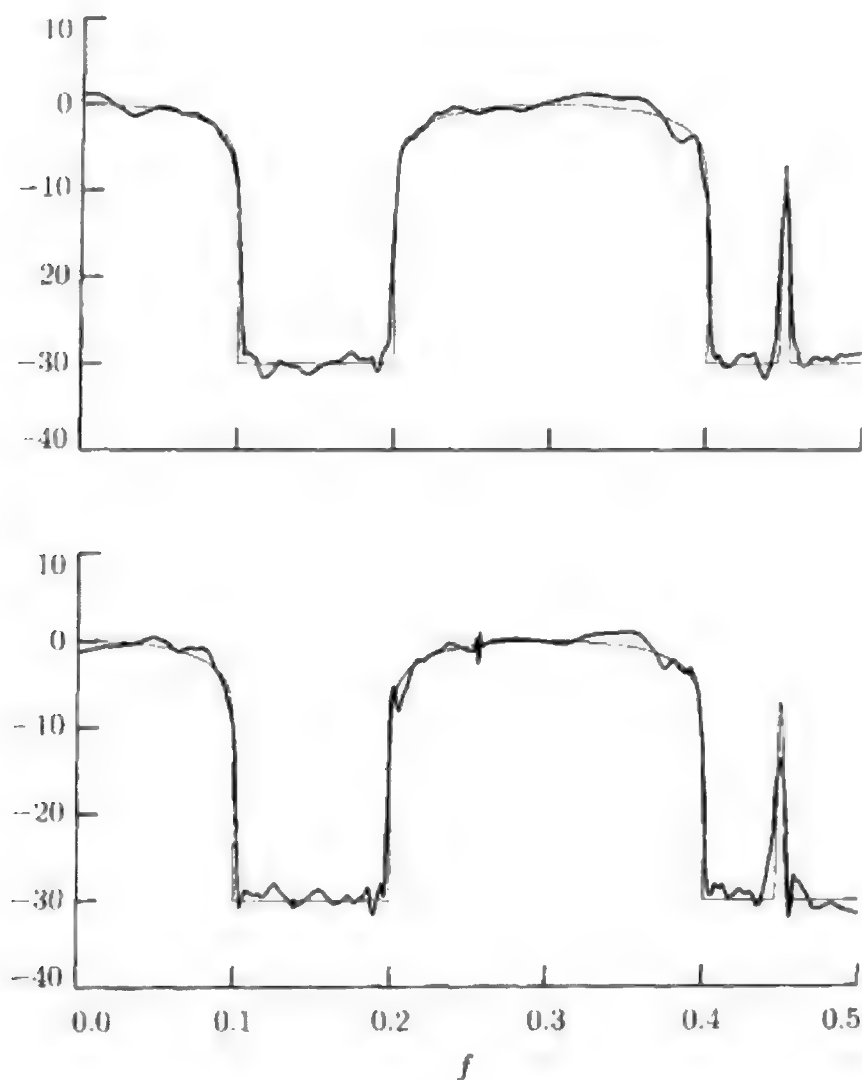


图 449 对于移动无线电通信过程的估计谱密度函数(粗曲线)和真谱密度函数(细曲线). 设计和参数参见图 447

面图形的细水平线表明了层独立的噪声方差 σ_η^2. 因为这个方差落在仅高于最低的两三层的层依赖方差 σ_j^2，多锥的相关性诱导出与下面对小波阈值的范例是一致的. 原因在于它实际上抑制了小尺度的“噪声尖峰”而留下了有用的大尺度小波系数没有削弱.

10.7 节的评论与扩展

[1]在多锥谱密度函数估计问题中，我们知道每一层的噪声小波系数的方差 σ_j^2. 因此，给出 $W_{j,t}$的实现 $w_{j,t}$，极小化

$$\sum_{t=0}^{N_j-1}(2\sigma_j^2 - w_{j,t}^2 + \delta_j^2)1_{[\delta_j^2,\infty)}(w_{j,t}^2) \tag{449}$$

作为 $\delta>0$ 的一个函数给我们一个对 j 层的系数的基于 Stein 无偏风险估计阈值 $\delta_j^{(s)}$(比较式(406b)).

[2]Johnstone and Silverman(1997)指出，在不知道方差的一般情况下，我们可以用在每个 j 层的中位数绝对偏差估计，即

$$\hat{\sigma}_{(\text{mad}),j} \equiv \frac{\text{median}\{|w_{j,0}|, |w_{j,1}|, \cdots, |w_{j,N_j-1}|\}}{0.6745},$$

然后在每一层用统一阈值

$$\hat{\delta}_j^{(u)} \equiv \sqrt{[2\hat{\sigma}^2_{(\text{mad}),j}\log(N)]}.$$

另一方面，我们可变为在每一层应用 Stein 无偏风险估计；即用式(405a)在一层一层的基础上获得 $\hat{\delta}_j^{(s)}$，作为 $\delta_j \geqslant 0$ 的函数使得

$$\sum_{t=0}^{N_j-1}(2\hat{\sigma}^2_{(\text{mad}),j} - w^2_{j,t} + \delta_j^2)1_{[\delta_j^2,\infty)}(w^2_{j,t})$$

极小. 如 Johnstone and Silverman(1997)所述，$\hat{\sigma}_{(\text{mad}),j}$ 只对小的 j 估计是适当的，其中有一个在给出层可考虑的数目，信号是稀疏的(即 $w_{j,t}$ 多是噪声).

[3]关于太阳地震数据的多锥谱估计的小波去噪估计的最新应用，可参见 Komm 等(1999).

10.8 小波系数的聚集和持续

在上节中，我们集中于对统计噪声小波系数的模型. 最近 Crouse 等(1998)着眼于统计信号离散小波变换系数的统计模型余下的问题. 为了激发他们的方法，他们指出对信号的统计模型应当考虑在观察的实际信号的经验离散小波系数的两个主要性质. 第一个性质就是聚集，也就是说，如果一个在 j 层内的特别的离散小波系数是大或者小的，则在同一层的伴随系数也可能是大或小的. 第二个性质是持续，也就是说大的或小的小波系数倾向于从不同层传播. 这些想法对心电图时间序列的 LA(8)离散小波系数由图 127 解释(这是一个对于把该序列作为一个无噪声信号想象的延伸，但是我们本节总是为教学起见这样做). 聚集特别体现在 $j=3$ 层的系数 $\mathbf{W}_3$. 当一个 R 波事件发生时，这里有两个或更大系数相互伴随. 持续性也与 R 波事件相关，可以从共同对准的大系数看出，如 $j=2$ 和 3 层. 合到一起，聚集和持续表明离散小波变换完全不能与实际信号完全无关，因此我们必须移到对统计信号的简单独立同分布模型上.

有了这些想法，Crouse 等(1998)发展了一个对模型小波系数独立性和非高斯性，围绕隐含马尔可夫模型(HMM)的新架构—参见 Rabiner(1989)对隐含马尔可夫模型的一个很好的回顾. 在离散小波变换中，对每一个信号离散小变换波系数隐含马尔可夫模型思想与一个隐含状态变量相关(特别注意状态变量不直接与原始时间序列或信号相关，而是与信号的变换有关). 为了符合信号离散小波变换系数的非高斯性，他们把每个系数的频率密度函数模型作为包含两个或更多

分量的一个混合高斯密度(图 410 右边图形描绘了两个分量混合密度的频率密度函数). 隐含的状态变量可以取状态值有限集合中的一个, 而状态值可以是相应于信号离散小波系数的一些统计, 即混合高斯的分量相应于频率密度函数. 尽管给出了它隐含的状态变量的值每一个信号离散小波系数是条件高斯的, 系数有一个完全(无条件的)非高斯密度(再次见图 410).

为了允许信号离散小波系数间的依赖, Crouse 等(1998)介绍了马尔可夫依赖性在基于对于小波和尺度系数的自然树结构的隐含状态变量. 为了了解树结构是怎样发生的, 假设我们计算对于信号的 J_0 层的部分离散小波变换系数, 产生向量 $\mathbf{W}_1$, …, $\mathbf{W}_{J_0}$ 和 $\mathbf{V}_{J_0}$. 我们能够绘出向量 $\mathbf{V}_{J_0}$ 和 $\mathbf{W}_{J_0}$, …, $\mathbf{W}_1$ 作为行堆在每一个上面, 如图 127, 这描绘了心电图序列的 $J_0=6$ 层部分离散小波变换(再次, 我们假设这是一个无噪声信号). 这里每一个系数向量已被循环平移, 使它可绘出逆物理意义的时间. 如果我们忽略了那幅图中的由 $\mathbf{V}_6$ 和 $\mathbf{W}_6$ 相关的很小的时间序列的平移而引起的复杂性, 那么"节点" $V_{6,2}$ 高于 $W_{6,3}$, 转而依次高于 $W_{5,3}$ 和 $W_{5,4}$. 这样数下去, $W_{5,3}$ 高于 $W_{4,3}$ 和 $W_{4,4}$, 而 $W_{5,4}$ 高于 $W_{4,5}$ 和 $W_{4,6}$, 依此类推. 换句话说, 在每一 J_0层系数下有一个树结构. 正式地, 我们可以说每一个在 $\mathbf{V}_{J_0}$ 中的 N_{J_0} 个尺度系数的每个一高于在 $\mathbf{W}_{J_0}$ 的一个系数, 而在 $\mathbf{W}_j$ 的每一个系数高于在 $\mathbf{W}_{j-1}$($j=J_0$, …, 2)中的两个系数. 因此这个结构就定义了 N_J 树. 由于一个状态变量与每一个小波或尺度系数有关. 伴随的是其状态变量可用 N_{J_0} 树的一个集表示出来.

这样就定义了树结构, Crouse 等(1998)对于树中状态变量的关系考虑了下面的树模型.

[1]在相同树中状态变量之间或不同树之间没有联系. 这认为离散小波变换状态变量(因此信号离散小波变换系数)作为独立的, 并且不计较聚集和持续.

[2]状态变量在 j 层间(同层通过树连接)有一阶马尔可夫链依赖(Rabiner, 1989), 但是由层到层之间是独立的(在同一树间不与下面层相关). 这组成了隐含马尔可夫链模型, 可考虑聚集但是没有持续.

[3]状态变量是垂直连接(一个树内向下层); 即有 N_{J_0} 树, 每一个由连接的状态变量组成. 这就是隐含马尔可夫树模型并且能提供聚集和持续.

Crouse 等(1998)强调对[2]和[3]的情况是建立在隐含的状态变量上的马尔可夫结构, 而不是直接建立在 $\mathbf{X}$ 的离散小波变换系数本身. 然后他们考虑统计信号加统计独立同分布高斯噪声的离散小波变换, 信号和噪声假设为独立的. 附加这种结构, 因为信号离散小波变换系数遵循隐含马尔可夫模型, 因此对 $\mathbf{X}$ 的离散小波变换也这么做. 为此, 注意给出隐含状态变量的值, 信号的离散小波系数是高斯的, 并且两个独立的高斯随机变量(信号加噪声)是高斯的, 方差是由方差的和给出. 因此信号离散小波变换模型诱导了对于 $\mathbf{X}$ 的离散小波变换的一个

平行的结构，我们可以用对观察的离散小波系数找出合适的隐含马尔可夫模型来估计信号(即那些对 $\mathbf{X}$). Crouse 等(1998)把信号离散小波变换系数的变量给定一个特别的状态变量，即观察的离散小波系数给出了同样的状态值，小于噪声的离散小波系数的方差(我们可以用$\hat{\sigma}^2_{(\text{mad})}$在等式(420)中的定义来估计后者). 关于这个估计方法更多的细节，参见 Crouse 等(1998).

10.9 小结

令 $\mathbf{X}$ 是一个 N 维向量，其分量为实值时间序列$\{X_t: t=0, \cdots, N-1\}$. 令 O_l 是 $\mathbf{O}$ 的第 l 个分量，系数向量来自于对 $\mathbf{X}$ 应用规范正交变换$\mathcal{O}$. 我们把 $\mathbf{X}$ 考虑为一个信号向量和噪声向量的和. 我们认为信号向量或是确定性(固定的)或统计的(随机的)，统计噪声向量总是具有零均值，但包含独立同分布的随机变量或非独立同分布的随机变量(改变独立同分布假设的变量更典型，取两种形式之一即随机变量是独立但是有不同的方差或者变量是相关的). 因此，我们认为 $\mathbf{X}$ 遵守下面四个模型中的一个：

[i] $\mathbf{X}=\mathbf{D}+\boldsymbol{\varepsilon}$，其中 $\mathbf{D}$ 为一个确定信号而ε为独立同分布噪声，所以应用规范正交变换给出

$$\mathbf{O}\equiv\mathcal{O}\mathbf{X}=\mathcal{O}\mathbf{D}+\mathcal{O}\boldsymbol{\varepsilon}\equiv\mathbf{d}+\mathbf{e}, \text{于是 } O_l=d_l+e_l;$$

[ii] $\mathbf{X}=\mathbf{D}+\boldsymbol{\eta}$. 其中 $\mathbf{D}$ 为一个确定信号而η为非独立同分布噪声，因此

$$\mathbf{O}\equiv\mathcal{O}\mathbf{X}=\mathcal{O}\mathbf{D}+\mathcal{O}\boldsymbol{\eta}\equiv\mathbf{d}+\mathbf{n}, \text{于是 } O_l=d_l+n_l;$$

[iii] $\mathbf{X}=\mathbf{C}+\boldsymbol{\varepsilon}$，其中 $\mathbf{C}$ 是一个统计信号而ε是独立同分布噪声(信号和噪声是相互独立的)，因此

$$\mathbf{O}\equiv\mathcal{O}\mathbf{X}=\mathcal{O}\mathbf{C}+\mathcal{O}\boldsymbol{\varepsilon}\equiv\mathbf{R}+\mathbf{e}, \text{于是 } O_l=R_l+e_l;$$

[iv] $\mathbf{X}=\mathbf{C}+\boldsymbol{\eta}$，其中 $\mathbf{C}$ 是一个统计信号，而η是非独立同分布噪声，因此

$$\mathbf{O}\equiv\mathcal{O}\mathbf{X}=\mathcal{O}\mathbf{C}+\mathcal{O}\boldsymbol{\eta}\equiv\mathbf{R}+\mathbf{n}, \text{于是 } O_l=R_l+n_l;$$

(典型的是信号和噪声相互独立，但是，例如在 Vidakovic(1998)研究的模型中这不是真的.)

假设在[iii]或[iv]中的感兴趣对于任意分量 l，规范正交变换给出 $E\{R_l\}=0$ (如果我们具体到离散小波变换，零均值假设不是完全严格的；如果随机信号 $\mathbf{C}$ 是 d 阶向后差分是平稳的随机过程而来的才可以保证满足，我们取滤波器长度 L 使 $L>2d$).

在输入独立同分布噪声情况下([i]和[iii]的情况)，噪声变换系数有相同的方差(无论ε是怎样分布的)，而且如果是高斯分布的则有相同的分布. 以符号形式，我们有

$$\Sigma_{\boldsymbol{\varepsilon}}=\sigma_\varepsilon^2 I_N \quad \text{和} \quad \Sigma_{\mathbf{e}}=\mathcal{O}\,\Sigma_{\boldsymbol{\varepsilon}}\,\mathcal{O}^{\mathrm{T}}=\sigma_\varepsilon^2 I_N, \quad \text{因此} \quad \operatorname{var}\{e_l\}=\operatorname{var}\{\varepsilon_l\}=\sigma_\varepsilon^2,$$

如果 $\varepsilon_l \stackrel{d}{=} \mathcal{N}(0, \sigma_\varepsilon^2)$，则 $e_l \stackrel{d}{=} \mathcal{N}(0, \sigma_\varepsilon^2)$. 在非独立同分布噪声情况下([ii]和[iv]情形)我们只能说

$$\Sigma_n = \mathcal{O}\,\Sigma_\eta\,\mathcal{O}^T;$$

然而，如果 $\eta_l \stackrel{d}{=} \mathcal{N}(0, \sigma_{\eta_l}^2)$那么 $n_l \stackrel{d}{=} \mathcal{N}(0, \sigma_{n_l}^2)$，但是不必有 $\sigma_{n_l}^2 = \sigma_{\eta_l}^2$，并且方差不依赖于下标 l.

我们可以用阈值或收缩方法进行信号估计，但是都需要如下步骤：

[1]计算变换系数 $\boldsymbol{O} \equiv \mathcal{O}\boldsymbol{X}$.

[2]定义 $\boldsymbol{O}^{(\cdot)}$ 向量，它的第 l 个分量是 $O_l^{(\cdot)}$，它是 O_l 进行了阈值或收缩后的分量，其中一般的角标“(·)”被替换为如“(t)”或“(shr)”以表示阈值或收缩(或更具体的如“(ht)”表示硬阈值).

[3]通过 $\mathcal{O}^T\boldsymbol{O}^{(\cdot)}$ 估计信号. 这里我们不区分确定性信号估计和统计信号的实现估计.

我们重点强调这个算法实现阈值或收缩的分量，即逐项进行. 阈值对于情况[i]和[ii]是方便的，即信号是确定性的. 法则包括硬、中和软阈值(分别是式(399),(400b)和(400a))，所有的由一个简单的阈值参数 $\delta>0$ 确定，而严格阈值(式(400c))由两个参数确定. 对于 O_l 的阈值层 δ 必须定义，并且依赖于噪声变换系数的统计性质的一个合适的定义. 如果这些系数是独立同分布，那么，用于所有的 O_l 是一个简单阈值. 对独立同分布高斯噪声，一个通常的阈值选择是统一阈值

$$\delta^{(u)} \equiv \sqrt{[2\sigma_\varepsilon^2 \log(N)]}$$

(这是式(400d)). 选择阈值水平的另外方法是通过使用交叉有效的极小化准则(式(422)和(402))，并且通过基于 Stein 的无偏风险估计(SURE)的一个方法，其中给出的随机变量的 O_l 的观察值 o_l 和假设的软阈值，我们必须设 $\delta^{(s)}$ 为使得下面式子极小的 δ：

$$\sum_{l=0}^{N-1}(2\sigma_\varepsilon^2 - o_l^2 + \delta^2)1_{[\delta^2,\infty)}(o_l^2)$$

(这是式(406a)，具有共同方差 σ_0^2 设为 σ_ε^2). 如果噪声变换系数是由非独立同分布，输入噪声 η，则阈值水平需要适当的变化. 例如，当系数分为组 C_j，在组内为独立同分布，有共同方差 σ_j^2，对于系数合适的阈值，在 j 组内可通过层依赖“统一”阈值来确定

$$\delta_j^{(u)} \equiv \sqrt{[2\sigma_j^2 \log(N)]}$$

(这就是式(406c))或者通过 Stein 无偏风险估计方法确定 $\delta_j^{(s)}$，作为使得下式极小的 δ：

$$\sum_{l\in C_j}(2\sigma_j^2 - o_{j,l}^2 + \delta^2)1_{[\delta^2,\infty)}(o_{j,l}^2)$$

(这是式(406b)).

收缩适合于情况[iii]和[iv]，即信号是统计模型的. 最简单的收缩法则是对观察的用信号方差 $\sigma_{R_l}^2$ 和总方差 $\sigma_{R_l}^2+\sigma_{n_l}^2$ 的比来重新尺度的(参见 10.3 节). 对于非线性收缩函数最著名的陈述包括信号条件均值 $E\{R_l \mid O_l=o_l\}$，给出观察(式(410b)). 例如，稀疏信号模型

$$R_l \stackrel{\mathrm{d}}{=} (1-\mathcal{I}_l)\,\mathcal{N}(0,\sigma_{G_l}^2),\quad n_l \stackrel{\mathrm{d}}{=} \mathcal{N}(0,\sigma_{n_l}^2)$$

导致条件均值

$$E\{R_l \mid O_l=o_l\}=\frac{b_l}{1+c_l}o_l,$$

其中 p_l 是二值随机变量 $\mathcal{I}_l$ 是 1 的概率，我们有

$$b_l=\frac{\sigma_{G_l}^2}{\sigma_{G_l}^2+\sigma_{n_l}^2},\quad c_l=\frac{p_l\sqrt{(\sigma_{G_l}^2+\sigma_{n_l}^2)}}{(1-p_l)\sigma_{n_l}}\mathrm{e}^{-o_l^2 b_l/(2\sigma_{n_l}^2)}$$

(参见式(411b)，(410e)，(411c)，(411a)和(411d)). 我们也可用条件均值代替(式(411e)). 收缩函数建立在信号变换系数 $\boldsymbol{R}$ 和噪声变换系数($\boldsymbol{e}$ 或 $\boldsymbol{n}$)的统计性质的基础上. 收缩法则详细定义了收缩函数的全局形状和系数收缩成很小值的区域. 收缩法则的确定因此等价于详细化阈值和阈值函数(比较图 412 和 399). 贝叶斯方法是通过贝叶斯法则定义收缩函数. 一些在文献中定义的后验均值和后验中值贝叶斯法则，原来与收缩函数中研究的条件均值和中值相同.

关于信号估计的这些结果对于一般的规范正交变换系数 O_l 是有效的，但它们能够有准备地特殊化并应用于离散小波变换情况. 阈值和收缩算法由下列的步骤组成：

[1]计算 J_0 层的部分离散小波变换分解获得 $\mathbf{W}_1$，…，$\mathbf{W}_{J_0}$ 和 $\mathbf{V}_{J_0}$.

[2]对每个小波系数 $W_{j,t}$，确定

(a)阈值水平用预先选的阈值函数或

(b)应用收缩法则.

[3]对于 $j=1$，…，J_0 和 $t=0$，…，N_j-1，令 $W_{j,t}^{(\cdot)}$ 是小波系数 $W_{j,t}$ 的阈值或收缩后的值，其中一般角标(·)可替代为，例如，"(t)"或"(shr)"，以确切表明是怎样处理 $W_{j,t}$ 的.

[4]对 $\mathbf{W}_1^{(\cdot)}$，…，$\mathbf{W}_{J_0}^{(\cdot)}$ 和 $\mathbf{V}_{J_0}$ 用逆变换估计信号.

对于输入独立同分布高斯噪 ε(参见 10.5 节)，简单阈值水平能够通过

$$\hat{\delta}^{(u)}=\sqrt{[2\,\hat{\sigma}_{(\mathrm{mad})}^2\log(N)]}$$

来估计，或者，给出随机变量 $W_{j,t}$ 的实现 $w_{j,t}$，通过基于 Stein 无偏风险估计作为 δ 的函数的估计 $\hat{\delta}^{(s)}$ 使得下面的式子极小：

$$\sum_{j=1}^{J_0}\sum_{t=0}^{N_j-1}(2\,\hat{\sigma}_{(\mathrm{mad})}^2-w_{j,t}^2+\delta^2)1_{[\delta^2,\infty)}(w_{j,t}^2),$$

其中

$$\hat{\sigma}_{(\mathrm{mad})} \equiv \frac{\mathrm{median}\{|w_{1,0}|,|w_{1,1}|,\cdots,|w_{1,\frac{N}{2}-1}|\}}{0.6745}$$

(上面分别是式(421)和(420)).

收缩法则能够容易地用于给出的统计模型结构. 当信号和噪声变量组成 $W_{j,t}$ 假设具有谱概率密度 j 层或位置 t 不变的参数，对于稀疏信号模型的收缩法则取式(425)的形式，即

$$w_{j,t}^{(\mathrm{shr})} = \frac{\hat{b}}{1+\hat{c}} w_{j,t},$$

其中 $\hat{b}$ 和 $\hat{c}$ 是参数 b 和 c 的估计(图 425 表示由 $w_{j,t}^{(\mathrm{shr})}$ 形成的信号估计的例子).

当噪声小波系数的频谱概率密度随层 j 的变化而发生形状变化时(这对 10.6 节讨论的基于周期图谱估计是正确的)，阈值的公式化就成了问题. 相关的一但是高斯的一噪声小波系数的谱估计算法在 10.7 节中提出，其中每个变换层 j 的方差 σ_j^2 已知. 因为这相当意外的先验知识，很容易用层依赖的统一阈值式(406c)实行阈值，在我们在 M 个非负的傅里叶频率采用了对数多锥估计值之后：

$$\delta_j^{(\mathrm{u})} \equiv \sqrt{[2\sigma_j^2 \log(M)]}$$

(这是式(444a)). 更惊奇地，应用层独立统一阈值式(443)也产生了好的结果，但回顾起来，这是有道理的：层独立统一阈值与小波阈值作用是相一致的，作用在于抑制了小尺度“噪声尖峰”而留下了大尺度系数未减小.

最后，在 10.8 节中，我们总结了 Crouse 等(1998)最近的工作，他指出对现实世界的信号的离散小波变换系数易于有在一层内大值聚集和层之间的持续性. 他们提出通过用隐含的马尔可夫树来模型化离散小波变换信号系数来保持聚集和持续性质，在其中状态变量与每一个信号离散小波变换系数相关，而且状态变量与树结构相关. 这样的方法为更实际的模型打开了一条通路.

10.10 练习

[10.1] 通过计算和绘出对于信号 $\mathbf{D}_1$，$\mathbf{D}_2$ 和 $\mathbf{D}_3$ 基于 LA(8)离散小波变换的规范的部分能量序列，证明在图 396 的右列的虚曲线是正确的. 对每一个信号用哈尔和 D(4)离散小波变换计算相同的规范的部分能量序列. 后两个离散小波变换的规范的部分能量序列相比于LA(8)情况及对恒等变换和正交离散傅里叶变换的规范部分能量序列如何?

[10.2] 假设 $\mathbf{X} \equiv \mathbf{D} + \boldsymbol{\varepsilon}$ 是长度为 $N=2^J$ 的时间序列，对某些 $J \geq 4$ 的整数，其中 $\mathbf{D}$ 是确定性信号，$\boldsymbol{\varepsilon}$ 是 N 维独立同分布高斯随机变量，均值为零，方差为 σ_ε^2. 如果我们基于哈尔离散小波变换的硬阈值估计 $\hat{\mathbf{D}}^{(\mathrm{ht})}$ 来估计 $\mathbf{D}$(像

4.1 节定义的)结合统一阈值 $\delta^{(u)}\equiv\sqrt{[2\sigma_\epsilon^2\log(N)]}$ (Donoho and Johnstone, 1994), 证明

(a)如果 $\boldsymbol{D}$ 是由(401a)定义的周期化构造扩展 16 点序列，那么，对一些小的 $\alpha>0$，我们有 $\boldsymbol{P}\left[\frac{1}{N}\|\hat{\boldsymbol{D}}^{(ht)}-\boldsymbol{D}\|^2<\alpha\right]\to 0$，当 $J\to\infty$ 时，即 $\hat{\boldsymbol{D}}^{(ht)}$ 不是 $\boldsymbol{D}$ 的一个一致的估计.

(b)另一方面，如果 $\boldsymbol{D}$ 由等式(401b)构造，那么，对任一 $\alpha>0$，我们有 $\boldsymbol{P}\left[\frac{1}{N}\|\hat{\boldsymbol{D}}^{(ht)}-\boldsymbol{D}\|^2<\alpha\right]\to 1$，当 $J\to\infty$ 时，即 $\hat{\boldsymbol{D}}^{(ht)}$ 是 $\boldsymbol{D}$ 的一个一致的估计.

[10.3] 证明，对软阈值，在式(404b)基于 Stein 无偏风险估计的函数 $\mathcal{R}(\cdot,\cdot,\cdot)$ 能够写成替换的形式

$$\mathcal{R}(\sigma_{n_l},O_l,\delta)=\sigma_{n_l}^2-2\sigma_{n_l}^2 1_{[0,\delta^2)}(O_l^2)+\min\{O_l^2,\delta^2\}.$$

[10.4] Gao(1998)提出对软、中、硬阈值的替换阈值法则称为非负绞刑的替换. 这个映射定义为

$$O_l^{(g)}=O_l\left(1-\frac{\delta^2}{O_l^2}\right)_+=\begin{cases}0, & |O_l|\leqslant\delta,\\ O_l-\frac{\delta^2}{O_l}, & \text{其他}.\end{cases}$$

证明，如果我们使用 Stein 无偏风险估计的阈值法则，在式(404b)中定义的函数 $\mathcal{R}(\sigma_{n_l}, O_l, \delta)$替换为

$$\mathcal{R}(\sigma_{n_l},O_l,\delta)=O_l^2-\sigma_{n_l}^2+\left(2\sigma_{n_l}^2-O_l^2+\frac{\delta^4+2\sigma_{n_l}^2\delta^2}{O_l^2}\right)1_{[\delta^2,\infty)}(O_l^2).$$

(能够证明，非负绞刑使得 $A^{(\delta)}(\cdot)$弱可微，如使 Stein 无偏风险估计所需要的.)

[10.5] 假设 R_l 和 n_l是独立的如等式(410d)(410e)分布的. 证明，在式(410b)定义的条件均值可写成式(410f)形式. 提示：根据在分母中涉及的两个积分用分步积分来表达(410f)的分子，并且完成被积函数的平方以得到可识形式的积分(即高斯的谱密度函数).

[10.6] 验证式(414b).

[10.7] 用 $J_0=6$ 层 LA(8)部分离散小波变换结合统一阈值，及硬、中和软阈值法则对心电图数据应用在 10.5 节中描述的基于小波阈值过程，但是把保留尺度系数不变替代为设置所有的尺度系数到零. 绘出三个信号估计并解释在不同估计中任何可注意的不同之处(心电图时间序列可从本书的网页上获取).

[10.8] 如 10.6 节中讨论的，我们对(重新定心)的对数周期图 $\boldsymbol{Y}^{(p)}$ 的部分离散小波变换应用基于小波阈值来获得对数谱估计 $\hat{\boldsymbol{D}}^{(t)}$. 通过构造，对于 $k=1,\cdots,M-1$，$\boldsymbol{Y}^{(p)}$ 的元素 k 和 $2M-k$ 是相同的，与频率 f_k 相关. 解释为什么相应的 $\hat{\boldsymbol{D}}^{(t)}$ 的元素即使两个都与频率 f_k 相关也不必相同.

[10.9] 通过用 D(4)和 LA(8)小波滤波器对 $M=2N=1\,024=2^{10}=2^J$ 和 $K=5$ 的情况，计算 $\sigma_1^2,\cdots,\sigma_J^2$ 来证明在式(441c)中陈述的层依赖变量的级数.

第 11 章

有限能量信号的小波分析

11.0 引言

连续时间小波变换正成为一个对连续时间"信号"进行多尺度描述的好工具，信号是定义在整个实轴上的一个有限能量函数．这个变换将一个信号与一个小波函数"伸长"方案（本质上是一个连续时间带通滤波器）有着本质的联系，从而得到这个信号的多分辨描述．本章，我们将总结连续时间小波变换的多分辨观点的重要概念和结果．我们的首要目的是阐述连续时间小波分析和前面第 4 章提供的离散小波分析之间的密切关系．为了阐述这一关系，我们首先在数学中讨论的关于小波的内积约定与工程中促成的滤波约定之间建立一种对应关系．为了简化，我们仅仅讨论定义在实域上的信号、尺度函数和小波函数．这里只讨论二进小波分析（基函数的伸缩因子为 2）进行分析．

本章给出的总结主要来自 Vetterli and Herley(1992)，Mallat(1989a，b，c) 以及 Daubechies(1988，1992)．为了进行顺利，我们首先假定所讨论性质的尺度函数和小波函数都存在．在 11.9 和 11.10 节中，我们讨论了采用 Daubechies (1988)方法来构造小波滤波器，其主要目的是为了理解消失距和正则性的概念．小波函数的构造的完整叙述由 Strichartz(1993)给出．

11.1 平移和伸缩

多分辨分析包括平移和伸缩两个运算，这两个运算都很容易对于定义在整个实轴上面的实值函数 $\gamma(\cdot)$来定义．对于单位平移运算，或者叫做延时，我们仍然采用记号$\mathcal{T}$（像 3.4 节的式(52a)中，其中它表示离散单位（循环）延时矩阵）．$\mathcal{T}$算子作用于函数 $\gamma(\cdot)$得到一个新的函数，我们表示为 $\gamma_{0,1}(\cdot)$．$\gamma_{0,1}(\cdot)$在给定时刻 t 的值为

$$\gamma_{0,1}(t) \equiv \gamma(t-1), -\infty < t < \infty.$$

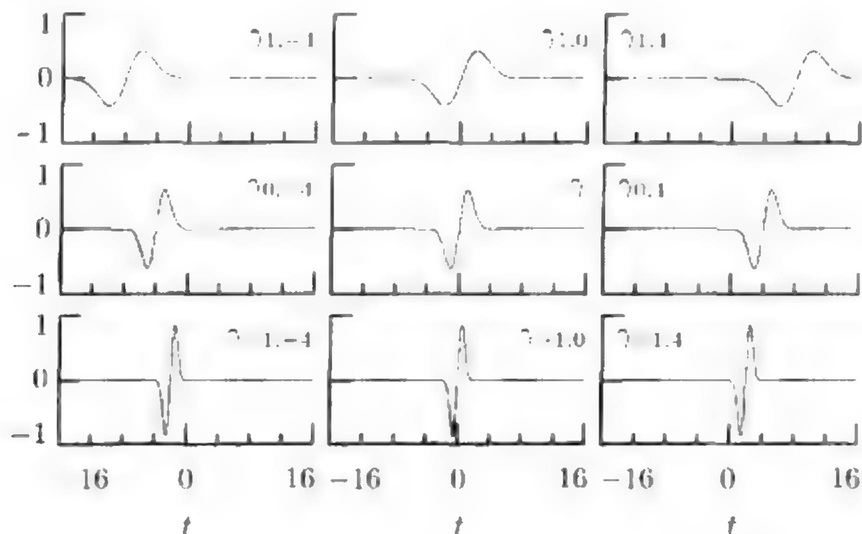


图 458 由 $\gamma(t)=t\exp(-t^2/2)$定义的函数的平移和伸缩

T作用于函数 $\gamma_{0,1}(\cdot)$得到函数 $\gamma_{0,2}(\cdot)$，它在时刻 t 的值是

$$\gamma_{0,2}(t)\equiv\gamma(t-2),\ -\infty<t<\infty;$$

类似地，我们能定义 $\gamma_{0,k}(t)$：

$$\gamma_{0,k}(t)\equiv\gamma(t-k),\ -\infty<t<\infty.$$

这个定义对所有 $k\in\mathbf{Z}$ 均成立($\mathbf{Z}$ 是所有整数的集合)．特别地，当 $k=-1$ 时，我们定义单位平移逆算子 T^{-1}，使 T^{-1} 作用于函数 $\gamma_{0,k}(\cdot)$上得到函数 $\gamma_{0,k-1}(\cdot)$ ($\gamma_{0,0}(\cdot)$定义是 $\gamma(\cdot)$)．

作为例子，如图 458 中间那一排图像，从左至右表示，对函数 $\gamma(t)=t\exp(-t^2/2)$的 $\gamma_{0,-4}(\cdot)$，$\gamma(\cdot)$和 $\gamma_{0,4}(\cdot)$．注意 $k>0$ 时，平移算子将原图像向右移动(相对于零时刻延时)，相反，$k<0$ 时，平移算子将原图像向左移(相对于零时刻提前)．

以类似的方式，我们能定义一个二进伸缩(尺度)算子 K．算子 K 作用于函数 $\gamma(\cdot)$上得到一个新的函数表示为 $\gamma_{1,0}(\cdot)$．在 t 时刻 $\gamma_{1,0}(\cdot)$的值为

$$\gamma_{1,0}(t)\equiv\frac{\gamma\left(\frac{t}{2}\right)}{\sqrt{2}},\ -\infty<t<\infty.$$

之所以包括因子$\frac{1}{\sqrt{2}}$是为了保持其范数不变．如果 $\gamma(\cdot)$是 $L^2(\mathbf{R})$(定义在 $\mathbf{R}$ 上的平方可积函数的集合)的元素，其平方范数为

$$\|\gamma(\cdot)\|^2\equiv\int_{-\infty}^{\infty}\gamma^2(t)\mathrm{d}t<\infty,$$

那么 $\|\gamma_{1,0}(\cdot)\|^2=\|\gamma(\cdot)\|^2$ (即伸缩算子保持范数，事实上平移算子也保持范数)．算子 K 作用于 $\gamma_{1,0}(\cdot)$得到函数 $\gamma_{2,0}(\cdot)$，在 t 时刻的值是

$$\gamma_{2,0}(t)\equiv\frac{\gamma\left(\frac{t}{4}\right)}{2},\quad -\infty<t<\infty;$$

类似地，我们能够定义 $\gamma_{j,0}(\cdot)$ 为

$$\gamma_{j,0}(\cdot)\equiv\frac{\gamma\left(\frac{t}{2^j}\right)}{\sqrt{2^j}},\quad -\infty<t<\infty.$$

同样这个定义对所有的 $j\in\mathbf{Z}$ 均成立，且选择 $j=-1$ 时，定义其为逆伸缩算子 $\mathcal{K}^{-1}$，算子 $\mathcal{K}^{-1}$ 作用在函数 $\gamma_{j,0}(\cdot)$ 上得到 $\gamma_{j-1,0}(\cdot)$.

作为二进伸缩的例子：图 458 中的中间一列，从上到下依次显示为 $\gamma_{1,0}(\cdot)$，$\gamma(\cdot)$ 和 $\gamma_{-1,0}(\cdot)$ 的图像. 注意 $j>0$ 时，函数图像直观上被扩展，$j<0$ 时相反.

最后，我们用 $(j,\ k)$ 表示对函数 $\gamma(\cdot)$ 平移继而伸缩所得到的新函数，定义为

$$\gamma_{j,k}(t)\equiv\frac{\gamma\left(\frac{t}{2^j}-k\right)}{\sqrt{2^j}},\quad -\infty<t<\infty.$$

在图 458 的左右两边上面的图像显示 $\gamma_{1,-4}(\cdot)$ 和 $\gamma_{1,4}(\cdot)$，而相应下面的显示 $\gamma_{-1,-4}(\cdot)$ 和 $\gamma_{-1,4}(\cdot)$. 注意 $\gamma_{j,k}(\cdot)$ 是 $\gamma_{0,k}(\cdot)$ 的第 j 次伸缩，但是一般 $\gamma_{j,k}(\cdot)$ 并不是 $\gamma_{j,0}(\cdot)$ 的第 k 次平移(即我们使用平移和伸缩算子在顺序上有很大关系，这些算子不是可交换的).

11.2 尺度函数和逼近空间

小波变换最初的发展是用作对一个有限能量信号 $x(\cdot)$ 进行分解和合成的工具，其中有限能量条件是

$$\|x\|^2\equiv\int_{-\infty}^{\infty}x^2(t)\mathrm{d}t<\infty$$

($x(\cdot)$ 仅限于实值)，即 $x(\cdot)$ 属于 $L^2(\mathbf{R})$. 令 $\phi(\cdot)\in L^2(\mathbf{R})$，使得它的整数平移 $\{\phi_{0,k}(\cdot):k\in\mathbf{Z}\}$ 形成某个闭子空间 $V_0\subset L^2(\mathbf{R})$ 的一个规范正交基，其中

$$V_0\equiv\overline{\mathrm{span}\{\phi_{0,k}(\cdot):k\in\mathbf{Z}\}}.$$

这里记号 span$\{S\}$ 表示线性张成，它是 S 中的元素的线性组合所形成的子空间，而记号 $\overline{\mathrm{span}\{S\}}$ 则表示由 S 生成的闭子空间(span$\{S\}$ 与 $\overline{\mathrm{span}\{S\}}$ 的区别在于后者包括线性组合的极限). 因此，我们有

$$\int_{-\infty}^{\infty}\phi_{0,k}(t)\phi_{0,m}(t)\mathrm{d}t=\int_{-\infty}^{\infty}\phi(t-k)\phi(t-m)\mathrm{d}t=\begin{cases}1, & k=m;\\ 0, & \text{其他}.\end{cases}\tag{459}$$

函数 $\phi(\cdot)$ 称为*尺度函数*，而空间 V_0 称为具有单位尺度的*逼近空间*. 我们可以通过 V_0 的基函数为背景投影 $x(\cdot)$ 在 V_0 中来分析任意函数 $x(\cdot)$. 如果我们定义

$$\langle x(\cdot),y(\cdot)\rangle\equiv\int_{-\infty}^{\infty}x(t)y(t)\mathrm{d}t$$

表示两个有限能量实值信号 $x(\cdot)$ 与 $y(\cdot)$ 的内积，$x(\cdot)$ 的分析产生

$$v_{0,k} \equiv \langle x(\cdot), \phi_{0,k}(\cdot) \rangle = \int_{-\infty}^{\infty} x(t)\phi_{0,k}(t)\mathrm{d}t = \int_{-\infty}^{\infty} x(t)\phi(t-k)\mathrm{d}t, \tag{460a}$$

其中 $v_{0,k}$ 称为函数 $x(\cdot)$ 的尺度系数. 如果 $x(\cdot) \in V_0$，我们可以由它的尺度系数来综合 $x(\cdot)$

$$x(t) = \sum_{k=-\infty}^{\infty} v_{0,k}\phi_{0,k}(t) = \sum_{k=-\infty}^{\infty} v_{0,k}\phi(t-k). \tag{460b}$$

如果 $x(\cdot) \notin V_0$，上面的无限求和是通过将 $x(\cdot)$ 在 V_0 上的投影来对 $x(\cdot)$ 的逼近.

作为一个简单的例子，考虑哈尔尺度函数，即

$$\phi^{(\mathrm{H})}(t) \equiv \begin{cases} 1, & -1 < t \leqslant 0; \\ 0, & \text{其他} \end{cases} \tag{460c}$$

（参见后面的评论与扩展的第[1]条）. 因此我们有

$$\phi_{0,k}^{(\mathrm{H})}(t) = \phi^{(\mathrm{H})}(t-k) = \begin{cases} 1, & k-1 < t \leqslant k; \\ 0, & \text{其他}. \end{cases}$$

$\phi_{0,k}^{(\mathrm{H})}(\cdot)$ 的规范正交性是容易建立的. 哈尔逼近空间 $V_0^{(\mathrm{H})}$ 由所有 $L^2(\mathbf{R})$ 中，在每个区间 $k-1<t\leqslant k (k\in \mathbf{Z})$ 上为常数定义的函数 $x(\cdot)$ 组成，即对某个实值常数 $v_{0,k}$，我们有 $x(t)=v_{0,k}$，对 $k-1<t\leqslant k$ 成立，其中 $v_{0,k}$ 能够由式(460a)得到. 图 461 的中间一行从左到右分别为 $\phi_{0,-1}^{(\mathrm{H})}(\cdot)$，$\phi^{(\mathrm{H})}(\cdot)$ 和 $\phi_{0,1}^{(\mathrm{H})}(\cdot)$，以及一个包含在 $V_0^{(\mathrm{H})}$ 中的函数的例子.

对式(459)中变量的一个改变如下：

$$\int_{-\infty}^{\infty} \phi_{0,k}(t)\phi_{0,m}(t)\mathrm{d}t = \int_{-\infty}^{\infty} \frac{\phi_{0,k}\left(\frac{t}{2^j}\right)}{\sqrt{2^j}} \frac{\phi_{0,m}\left(\frac{t}{2^j}\right)}{\sqrt{2^j}}\mathrm{d}t = \int_{-\infty}^{\infty} \phi_{j,k}(t)\phi_{j,m}(t)\mathrm{d}t,$$

其中

$$\phi_{j,k}(t) \equiv \frac{\phi_{0,k}\left(\frac{t}{2^j}\right)}{\sqrt{2^j}} = \frac{\phi\left(\frac{t}{2^j}-k\right)}{\sqrt{2^j}}.$$

因此，给定 $\{\phi_{0,k}(\cdot) : k\in \mathbf{Z}\}$ 构成空间 V_0 的规范正交基，由此得到 $\{\phi_{j,k}(\cdot) : k\in \mathbf{Z}\}$ 构成空间 V_j 的一个规范正交基，其中

$$V_j \equiv \overline{\mathrm{span}\{\phi_{j,k}(\cdot) : k \in \mathbf{Z}\}}.$$

尺度为 2^j 的逼近空间 V_j，因此是由 $\phi_{j,k}(\cdot)$，$k\in \mathbf{Z}$，所张成的闭的线性空间.

现在假设下面的条件成立：

$$\cdots \subset V_3 \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset \cdots.$$

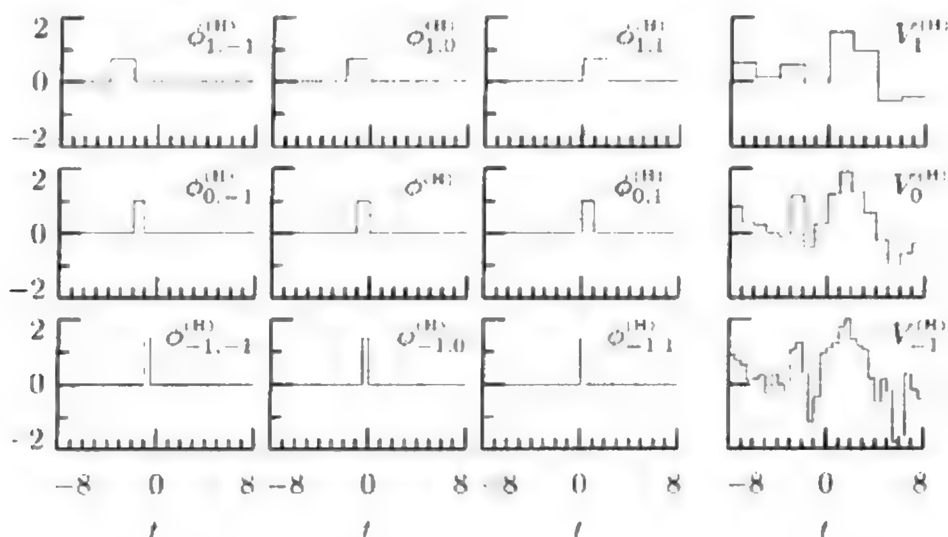


图 461 哈尔尺度函数 $\phi^{(H)}(\cdot)$ 和相应的逼近空间．中间行的前面三个图展示哈尔逼近空间 $V_0^{(H)}$ 的三个基函数，即从左到右，$\phi_{0,-1}^{(H)}(\cdot)$，$\phi^{(H)}(\cdot)$ 和 $\phi_{0,1}^{(H)}(\cdot)$．这一行的最右边是包含在空间 $V_0^{(H)}$ 内的一个函数的例子．最上面和最下面的两行分别表示哈尔逼近空间 $V_1^{(H)}$（比 $V_0^{(H)}$ 更粗糙的空间）和 $V_{-1}^{(H)}$（比 $V_0^{(H)}$ 更好的逼近空间）．最右边的一列可视为 $L^2(\mathbf{R})$ 内一个信号的相应于尺度 2，1 和 $\frac{1}{2}$ 的三个哈尔逼近（从上到下）

上述条件表明任一有限能量信号到 V_j 上的投影将给出逐次的逼近．如果我们想要这逐次的逼近对任何 $L^2(\mathbf{R})$ 上函数都实用，我们必须进而假设

$$\overline{\bigcup_{j\in\mathbf{Z}} V_j} = L^2(\mathbf{R}),\ \bigcap_{j\in\mathbf{Z}} V_j = \{0\}, \tag{461}$$

其中 0 表示零空间（即函数在整个实轴上恒为零）．这保证了对任意的 $y(\cdot)\in L^2(\mathbf{R})$，当 $j\to-\infty$ 时，它在 V_j 上的投影就等于 $y(\cdot)$．在嵌套子空间中，注意 $l<j$ 意味着 V_l 是比 V_j 更大的子空间（即 $V_j\subset V_l$）．这跟 Daubechies(1988，1992) 习惯一样，但是跟 Mallat(1989b) 的习惯相反．

作为例子，在图 461 的顶行显示，从左到右是尺度为 2 的哈尔逼近空间 $V_1^{(H)}$ 的三个基函数和包含在这个空间中的一个函数的例子（这三个基函数从左到右依次为：$\phi_{1,-1}^{(H)}(\cdot)$，$\phi_{1,0}^{(H)}(\cdot)$ 和 $\phi_{1,1}^{(H)}(\cdot)$）．底行显示尺度为 1/2 哈尔逼近空间 $V_{-1}^{(H)}$ 的相应的图．最右边一列，从上到下，对某个 $L^2(\mathbf{R})$ 的函数 $x(\cdot)$ 逐次地更好的逼近，其中逼近是通过向子空间 $V_1^{(H)}$，$V_0^{(H)}$ 和 $V_{-1}^{(H)}$ 投影得到的（参见式(462b)）．

本节的重点可以概括为以下定义：(Jawerth and Sweldens，1994；Daubechies，1988；Mallat，1989b)：*多分辨分析*（MRA）是定义在 $L^2(\mathbf{R})$ 上的一列闭子空间 V_j，$j\in\mathbf{Z}$ 使

[1]$V_j \subset V_{j-1}$，对所有 $j \in \mathbf{Z}$.

[2]$x(\cdot) \in V_0$当且仅当 $x_{0,k}(\cdot) \in V_0$，其中 $x_{0,k}(t) \equiv x(t-k)$，$k \in \mathbf{Z}$.

[3]$x(\cdot) \in V_0$当且仅当 $x_{j,0}(\cdot) \in V_j$，其中 $x_{j,0}(t) \equiv x\left(\frac{t}{2^j}\right)/\sqrt{2^j}$，$j \in \mathbf{Z}$.

[4]空间 V_j满足式(461)的两个条件.

[5]存在一个尺度函数 $\phi(\cdot) \in V_0$，使 $\{\phi_{0,k}(\cdot) : k \in \mathbf{Z}\}$ 形成 V_0的一个规范正交基，其中 $\phi_{0,k}(t) \equiv \phi(t-k)$.

子空间 V_j称为对于尺度 $\lambda_j = 2^j$的逼近空间.

11.2 节的评论与扩展

[1]我们在式(460c)中对于哈尔尺度函数不是文献中通常的那个尺度函数——通常 $\phi^{(H)}$取正的区间为(0，1]. 如下一节的讨论，哈尔尺度函数诱导出对于哈尔滤波器的定义. 使用对 $\phi^{(H)}(\cdot)$ 非标准定义的根本理由是保证哈尔尺度函数能够得到在 4.3 节引入的同样的哈尔尺度滤波器；它在式(463d)中表示. 这个非标准定义的需要仅仅反映内积运算与卷积在时间上面的不同的事实.

11.3 有限能量信号的逼近

像前面一样，假设 $\{V_j : j \in \mathbf{Z}\}$ 构成 $L^2(\mathbf{R})$的一个多分辨分析. 如像式(460b)和(460a)指出的，信号 $x(\cdot) \in V_0$能够写为

$$x(t) \equiv \sum_{k=-\infty}^{\infty} \langle x(\cdot), \phi_{0,k}(\cdot) \rangle \phi_{0,k}(t) = \sum_{k=-\infty}^{\infty} v_{0,k} \phi_{0,k}(t), \tag{462a}$$

其中 $v_{0,k} \equiv \langle x(\cdot), \phi_{0,k}(\cdot) \rangle$. 因为 $x(\cdot) \in V_0$，所以我们说 $x(\cdot)$为一个尺度为 $\lambda_0 = 2^0 = 1$ 的信号. 我们可以得到 $x(\cdot)$的一个逼近，$x(\cdot)$向子空间 $V_1 \subset V_0$ 上投影得到

$$s_1(t) \equiv \sum_{k=-\infty}^{\infty} \langle x(\cdot), \phi_{1,k}(\cdot) \rangle \phi_{1,k}(t) = \sum_{k=-\infty}^{\infty} v_{1,k} \phi_{1,k}(t), \tag{462b}$$

其中

$$v_{1,k} \equiv \langle x(\cdot), \phi_{1,k}(\cdot) \rangle = \int_{-\infty}^{\infty} x(t) \phi_{1,k}(t) \mathrm{d}t = \int_{-\infty}^{\infty} x(t) \frac{\phi\left(\frac{t}{2} - k\right)}{\sqrt{2}} \mathrm{d}t. \tag{462c}$$

因为 $s_1(\cdot) \in V_1$，所以信号 $s_1(\cdot)$具有尺度 $\lambda_1 = 2^1 = 2$，并且 $s_1(\cdot)$表示单位尺度信号 $x(\cdot)$的一个粗糙的逼近. $v_{1,k}$称为尺度为 2 的尺度系数.

通过将式(462a)代入式(462c)，在 $v_{0,k}$和 $v_{1,k}$之间有一个有趣的关系，我们能够得到

$$v_{1,k} = \int_{-\infty}^{\infty} \left[\sum_{m=-\infty}^{\infty} v_{0,m} \phi_{0,m}(t) \right] \frac{\phi\left(\frac{t}{2} - k\right)}{\sqrt{2}} \mathrm{d}t$$

$$= \sum_{m=-\infty}^{\infty} v_{0,m} \int_{-\infty}^{\infty} \phi(t-m) \frac{\phi\left(\frac{t}{2}-k\right)}{\sqrt{2}} \mathrm{d}t$$

$$= \sum_{m=-\infty}^{\infty} v_{0,m} \int_{-\infty}^{\infty} \phi(t-[m-2k]) \frac{\phi\left(\frac{t}{2}\right)}{\sqrt{2}} \mathrm{d}t$$

$$= \sum_{m=-\infty}^{\infty} \bar{g}_{m-2k} v_{0,m}, \tag{463a}$$

其中

$$\bar{g}_l \equiv \int_{-\infty}^{\infty} \phi(t-l) \frac{\phi\left(\frac{t}{2}\right)}{\sqrt{2}} \mathrm{d}t = \langle \phi_{0,l}(\cdot), \phi_{1,0}(\cdot) \rangle. \tag{463b}$$

如果我们定义 $g_l \equiv \bar{g}_{-l}$，并且在式(463a)中令 $l=2k-m$，则可以写成

$$v_{1,k} = \sum_{l=-\infty}^{\infty} g_l v_{0,2k-l}.$$

因此$\{v_{1,k}\}$相应于用$\{g_l\}$滤波$\{v_{0,k}\}$产生的输出(即因子 2 下抽样)的每个另外的值(滤波器$\{g_l\}$与我们在 4.3 节中引入的尺度滤波器相同).

例如，让我们计算哈尔尺度函数 $\phi^{(H)}(\cdot)$的$\{\bar{g}_l\}$. 因为

$$\phi^{(H)}(t-l) = \begin{cases} 1, & l-1 < t \leqslant l; \\ 0, & \text{其他}; \end{cases} \quad \phi^{(H)}\left(\frac{t}{2}\right) = \begin{cases} 1, & -2 < t \leqslant 0; \\ 0, & \text{其他}. \end{cases}$$

由式(463b)得出

$$\bar{g}_l = \frac{1}{\sqrt{2}} \int_{-2}^{0} \phi^{(H)}(t-l) \mathrm{d}t = \begin{cases} 1/\sqrt{2}, & l=0 \text{ 或 } -1; \\ 0, & \text{其他}. \end{cases} \tag{463c}$$

由于 $g_l \equiv \bar{g}_{-l}$，我们有

$$g_l = \begin{cases} 1/\sqrt{2}, & l=0 \text{ 或 } 1; \\ 0, & \text{其他}. \end{cases} \tag{463d}$$

这与式(75c)给出的哈尔尺度滤波器$\{g_l\}$完全一致.

因为 $s_1(\cdot) \in V_1$，所以我们可以通过向子空间 $V_2 \subset V_1$ 上投影得到 $s_1(\cdot)$的一个逼近，得到

$$s_2(t) \equiv \sum_{k=-\infty}^{\infty} v_{2,k} \phi_{2,k}(t), \text{而 } v_{2,k} \equiv \langle s_1(\cdot), \phi_{2,k}(\cdot) \rangle,$$

$s_2(\cdot)$也表示尺度 $\lambda_2 = 2^2 = 4$ 的 $x(\cdot)$的逼近. 使用式(462b)对 $s_1(\cdot)$的表示式，我们有

$$v_{2,k} = \left\langle \sum_{m=-\infty}^{\infty} v_{1,m} \phi_{1,m}(\cdot), \phi_{2,k}(\cdot) \right\rangle$$

$$= \sum_{m=-\infty}^{\infty} v_{1,m} \langle \phi_{1,m}(\cdot), \phi_{2,k}(\cdot) \rangle$$

$$= \sum_{m=-\infty}^{\infty} \bar{g}_{m-2k} v_{1,m} = \sum_{l=-\infty}^{\infty} g_l v_{1,2k-l}, \tag{463e}$$

其中我们使用下述的练习结论.

练习[464]　证明：

$$\langle \phi_{j,m}(\cdot), \phi_{j+1,k}(\cdot) \rangle = \langle \phi_{0,m-2k}(\cdot), \phi_{1,0}(\cdot) \rangle = \bar{g}_{m-2k}$$

对于所有的 j 成立，其中 $\bar{g}_l$ 在式(463b)中定义(令 $j=1$ 即得式(463e))．作为一个推论，还证明

$$\langle \phi_{j,m}(\cdot), \phi_{j-1,k}(\cdot) \rangle = \bar{g}_{k-2m} \tag{464a}$$

对于所有的 j 成立．◁

一般来说，给定 $s_{j-1}(\cdot) \in V_{j-1}$，我们可以用它到子空间 $V_j \subset V_{j-1}$ 上投影逼近 $s_{j-1}(\cdot)$，得到

$$s_j(t) \equiv \sum_{k=-\infty}^{\infty} v_{j,k} \phi_{j,k}(t),$$

其中

$$v_{j,k} \equiv \langle s_{j-1}(\cdot), \phi_{j,k}(\cdot) \rangle = \langle x(\cdot), \phi_{j,k}(\cdot) \rangle.$$

信号 $s_j(\cdot)$ 表示尺度 $\lambda_j = 2^j$ 的对信号 $x(\cdot)$ 的逼近．对于 $v_{j,k}$ 的递推计算表示式为

$$v_{j,k} = \sum_{m=-\infty}^{\infty} \bar{g}_{m-2k} v_{j-1,m} = \sum_{l=-\infty}^{\infty} g_l v_{j-1,2k-l}. \tag{464b}$$

11.3 节的评论与扩展

[1]假设我们有一时间序列 $\{x_t : t \in \mathbf{Z}\}$ 能够认为是信号 $x(\cdot)$ 的抽样，即 $x_t = x(t)$．由式(460a)我们知道每一个初始尺度序列的每一个元素 $v_{0,k}$ 都是 $x(\cdot)$ 与 $\phi_{0,k}(\cdot)$ 的内积．如果对 $x(\cdot)$ 的了解仅限于 $\mathbf{Z}$ 上的它的值，即 $\{x_t\}$，那么严格地说我们不能够得到 $x(\cdot)$ 的一个多分辨分析，因为我们没有足够的信息来构造 $v_{0,k}$．然而，我们可以使用黎曼(Riemann)和来得到近似的 $v_{0,k}$：

$$v_{0,k} = \int_{-\infty}^{\infty} x(t)\phi_{0,k}(t)\,\mathrm{d}t \approx \sum_{t=-\infty}^{\infty} x(t)\phi_{0,k}(t) = \sum_{t=-\infty}^{\infty} x_t \phi_{0,k}(t).$$

实际上，这个算法的初始值常常取 naïve 逼近：$v_{0,k} = x_k$．Jawerth and Sweldens (1994)指出其他采样过程，包括建议的求积公式，并给出参考文献．Abry and Flandrin(1994)也给出了初始化效果的例子．

11.4　尺度函数的两尺度关系

本节我们探讨一下所谓的对于尺度方程的两尺度关系．首先我们注意到：既然 $V_0 \subset V_{-1}$，$\phi(\cdot) \in V_0$，推出 $\phi(\cdot) \in V_{-1}$，因此我们可以写成

$$\phi(t) = \sum_{l=-\infty}^{\infty} \langle \phi(\cdot), \phi_{-1,l}(\cdot) \rangle \phi_{-1,l}(t) = \sum_{l=-\infty}^{\infty} \bar{g}_l \phi_{-1,l}(t),$$

其中用到了式(464a)以得到$\langle \phi(\cdot), \phi_{-1,l}(\cdot)\rangle = \bar{g}_l$(在使用这个式子中，注意$\phi(\cdot)\equiv \phi_{0,0}(\cdot)$). 因为$\phi_{-1,l}(t)=\sqrt{2}\phi(2t-l)$，从而得到两尺度关系：

$$\phi(t) = \sqrt{2}\sum_{l=-\infty}^{\infty} \bar{g}_l \phi(2t-l), \tag{465a}$$

这有时也称为*两尺度差分方程*.

下面我们研究两尺度关系的一些结论. 假设$\phi(\cdot)$在$a<t\leqslant b$之外的值为0，此时我们也称$\phi(\cdot)$具有$(a, b]$上的*有限支撑*. 我们可以表示这个现实为support $\{\phi(\cdot)\}=(a, b]$. 因为函数$\phi(2t-l)$的支撑区间为$\left(\frac{a+l}{2}, \frac{b+l}{2}\right]$，根据式(465a)得到$\bar{g}_l$只有有限项为非零. 因此，两尺度方程右边的支撑区间为$\left(\frac{a-(L-1)}{2}, \frac{b}{2}\right]$，由左右支撑区间相等知，我们必定有$(a, b]=\left(\frac{a-(L-1)}{2}, \frac{b}{2}\right]$，这推出$a=-(L-1)$，$b=0$，即$\phi(\cdot)$在区间$(-(L-1), 0]$之外为0. 值得一提的是，如果滤波器$\{g_l\}$在$l\neq 0, \cdots, L-1$时为0，那么$\{\bar{g}_l\}\equiv\{g_{-l}\}$在$l\neq -L-1, \cdots, 0$时为0. 所以长度为$L$的尺度滤波器$\{\bar{g}_l\}$对应于支撑区间$(-(L-1), 0]$上的尺度函数$\phi(\cdot)$. 前面我们已经看到了一个简单的例子：哈尔尺度滤波器的长度为$L=2$，式(460c)哈尔尺度函数的支撑区间为$(-1, 0]$.

我们可以利用傅里叶变换借助于$\{\bar{g}_l\}$得到尺度函数$\phi(\cdot)$的解. 令

$$\Phi(f)\equiv\int_{-\infty}^{\infty}\phi(t)e^{-i2\pi ft}dt$$

是函数$\phi(\cdot)$的傅里叶变换. 因为

$$\int_{-\infty}^{\infty}\phi(2t-l)e^{-i2\pi ft}dt = \frac{e^{-i2\pi\frac{f}{2}l}}{2}\int_{-\infty}^{\infty}\phi(t)e^{-i2\pi\frac{f}{2}t}dt = \frac{e^{-i2\pi\frac{f}{2}l}}{2}\Phi\left(\frac{f}{2}\right),$$

对式(465a)两边进行傅里叶变换得到

$$\Phi(f) = \sqrt{2}\sum_{l=-\infty}^{\infty}\bar{g}_l\frac{e^{-i2\pi\frac{f}{2}l}}{2}\Phi\left(\frac{f}{2}\right) = \frac{\Phi\left(\frac{f}{2}\right)}{\sqrt{2}}\sum_{l=-\infty}^{\infty}\bar{g}_l e^{-i2\pi\frac{f}{2}l} = \Phi\left(\frac{f}{2}\right)\frac{\bar{G}\left(\frac{f}{2}\right)}{\sqrt{2}}, \tag{465b}$$

其中$\bar{G}(\cdot)$是$\{\bar{g}_l\}$的傅里叶变换，即

$$\bar{G}(f)\equiv\sum_{l=-\infty}^{\infty}\bar{g}_l e^{-i2\pi fl}.$$

经过n次迭代后得到

$$\Phi(f) = \Phi\left(\frac{f}{2^n}\right)\prod_{m=1}^{n}\frac{\bar{G}\left(\frac{f}{2^m}\right)}{\sqrt{2}}. \tag{465c}$$

当 $n\to\infty$ 时，我们有

$$\Phi(f)=\Phi(0)\prod_{m=1}^{\infty}\frac{\bar{G}\left(\frac{f}{2^m}\right)}{\sqrt{2}},\qquad(466a)$$

由此我们可以通过 $\Phi(\cdot)$的傅里叶逆变换得到尺度函数.

如果在式(465b)中令 $f=0$，并且如果我们回忆 $\Phi(0)$刚好是$\phi(\cdot)$的积分，就得到

$$\int_{-\infty}^{\infty}\phi(t)\,\mathrm{d}t=\sqrt{2}\sum_{l=-\infty}^{\infty}\bar{g}_l\int_{-\infty}^{\infty}\phi(2t-l)\,\mathrm{d}t=\frac{1}{\sqrt{2}}\sum_{l=-\infty}^{\infty}\bar{g}_l\int_{-\infty}^{\infty}\phi(t)\,\mathrm{d}t.$$

在 $\phi(\cdot)$的积分不等于零的假设下，我们有

$$\sum_{l=-\infty}^{\infty}\bar{g}_l=\sqrt{2},\text{这推出 }\bar{G}(0)=\sum_{l=-\infty}^{\infty}\bar{g}_l=\sqrt{2}.\qquad(466b)$$

因为

$$\phi_{-1,0}\left(t-\frac{l}{2}\right)=\phi_{-1,0}\left(\frac{2t-l}{2}\right)=\sqrt{2}\phi(2t-l),$$

两尺度差分方程(465a)能够写成

$$\phi(t)=\sum_{l=-\infty}^{\infty}\bar{g}_l\phi_{-1,0}\left(t-\frac{l}{2}\right).$$

类似地，差分方程也可以写成

$$\phi_{1,0}(t)=\sum_{l=-\infty}^{\infty}\bar{g}_l\phi_{0,0}(t-l)\text{ 或 }\phi_{2,0}(t)=\sum_{l=-\infty}^{\infty}\bar{g}_l\phi_{1,0}(t-2l),$$

更一般地

$$\phi_{j+1,0}(t)=\sum_{l=-\infty}^{\infty}\bar{g}_l\phi_{j,0}(t-2^jl).\qquad(466c)$$

练习[466] 令 $\Phi_{j+1}(\cdot)$表示 $\phi_{j+1,0}(\cdot)$的傅里叶变换. 证明

$$\Phi_{j+1}(f)=\bar{G}(2^jf)\Phi_j(f).$$ ◁

例如，从 $j=2$ 开始，注意 $\Phi_0(\cdot)\equiv\Phi(\cdot)$，经过上面结论的三次应用后可得

$$\Phi_3(f)=\bar{G}(4f)\bar{G}(2f)\bar{G}(f)\Phi(f)=\bar{G}(4f)\bar{G}(2f)\bar{G}(f)\Phi(0)\prod_{m=1}^{\infty}\frac{\bar{G}\left(\frac{f}{2^m}\right)}{\sqrt{2}}.$$

$\{\phi_{0,k}(\cdot): k\in\mathbf{Z}\}$构成规范正交函数集的事实配置 $\bar{G}(\cdot)$的形式上的一个重要约束. 由于$\phi(\cdot)\leftrightarrow\Phi(\cdot)$推出

$$\int_{-\infty}^{\infty}\phi_{0,k}(t)\mathrm{e}^{-\mathrm{i}2\pi ft}\,\mathrm{d}t=\mathrm{e}^{-\mathrm{i}2\pi fk}\Phi(f),$$

帕塞瓦尔定理的两个函数形式(式 38b)告诉我们

$$\delta_{k,0}=\int_{-\infty}^{\infty}\phi_{0,k}(t)\phi(t)\,\mathrm{d}t=\int_{-\infty}^{\infty}|\Phi(f)|^2\mathrm{e}^{\mathrm{i}2\pi fk}\,\mathrm{d}f,$$

其中 $\delta_{k,0}$，如果 $k=0$ 时是 1 并且在其他情况下为 0. 因此，我们有

$$\delta_{k,0}=\sum_{m=-\infty}^{\infty}\int_{(2m-1)/2}^{(2m+1)/2}|\Phi(f)|^2\mathrm{e}^{\mathrm{i}2\pi fk}\,\mathrm{d}f=\sum_{m=-\infty}^{\infty}\int_{-1/2}^{1/2}|\Phi(f+m)|^2\mathrm{e}^{\mathrm{i}2\pi(f+m)k}\,\mathrm{d}f$$

$$=\int_{-1/2}^{1/2}\left(\sum_{m=-\infty}^{\infty}|\Phi(f+m)|^2\right)\mathrm{e}^{\mathrm{i}2\pi fk}\,\mathrm{d}f.$$

以上等式表明圆括号内求和的傅里叶逆变换是序列$\{\delta_{k,0}\}$，而$\{\delta_{k,0}\}$的傅里叶变换在任何频率下都等于 1，所以得到

$$\sum_{m=-\infty}^{\infty}|\Phi(f+m)|^2=1. \tag{467a}$$

使用式(465b)，我们有

$$\sum_{m=-\infty}^{\infty}|\Phi(f+m)|^2=\frac{1}{2}\sum_{m=-\infty}^{\infty}\left|\bar{G}\left(\frac{f}{2}+\frac{m}{2}\right)\right|^2\left|\Phi\left(\frac{f}{2}+\frac{m}{2}\right)\right|^2,$$

也就是说

$$\sum_{m=-\infty}^{\infty}\left|\bar{G}\left(f'+\frac{m}{2}\right)\right|^2\left|\Phi\left(f'+\frac{m}{2}\right)\right|^2=2,$$

其中 $f'=f/2$. 上面和式的偶数 m 指标部分可以写成

$$\sum_{n=-\infty}^{\infty}|\bar{G}(f'+n)|^2|\Phi(f'+n)|^2=|\bar{G}(f')|^2\sum_{n=-\infty}^{\infty}|\Phi(f'+n)|^2=|\bar{G}(f')|^2,$$

其中我们利用了式(467a)和 $\bar{G}(\cdot)$的周期为 1 的周期性(即 $\bar{G}(f'+n)=\bar{G}(f')$，对所有的整数 n 均成立). 类似地，和式的奇数 m 指标部分可以写成

$$\sum_{n=-\infty}^{\infty}\left|\bar{G}\left(f'+n+\frac{1}{2}\right)\right|^2\left|\Phi\left(f'+n+\frac{1}{2}\right)\right|^2$$

$$=\left|\bar{G}\left(f'+\frac{1}{2}\right)\right|^2\sum_{n=-\infty}^{\infty}\left|\Phi\left(f'+\frac{1}{2}+n\right)\right|^2=\left|\bar{G}\left(f'+\frac{1}{2}\right)\right|^2.$$

联合得到约束

$$|\bar{G}(f)|^2+\left|\bar{G}\left(f+\frac{1}{2}\right)\right|^2=2. \tag{467b}$$

特别地，如果在式中令 $f=0$，就得到

$$|\bar{G}(0)|^2+\left|\bar{G}\left(\frac{1}{2}\right)\right|^2=2;$$

另一方面，式(466b)告诉我们 $\bar{G}(0)=\sqrt{2}$，从而有结论

$$\bar{G}\left(\frac{1}{2}\right)=0, \tag{468a}$$

该结论将在 11.6 节证明是有用的.

该结论同时让我们能够确定 $\phi(\cdot)$规范化. 首先回顾式(467a)，我们可以写成

$$|\Phi(f)|^2+\sum_{\substack{m=-\infty\\m\neq 0}}^{\infty}|\Phi(f+m)|^2=1,\text{因此}\,|\Phi(0)|^2+\sum_{\substack{m=-\infty\\m\neq 0}}^{\infty}|\Phi(m)|^2=1,$$

当 $f=0$ 时；然而，由下面，我们看到对最后和的所有项都有 $\Phi(m)=0$.

练习[468a] 令 m 为任一非零整数. 用记号 $m=2^l[2j+1]$，对某个 $l\geqslant 0$ 和 $j\in\mathbf{Z}$ 并且用式(465c)证明

$$\Phi(m)=\Phi\left(\frac{2j+1}{2}\right)\prod_{k=1}^{l+1}\frac{\bar{G}(2^{l-k}[2j+1])}{\sqrt{2}}.$$

并由此讨论 $\Phi(m)=0$. ◁

我们能够得出结论 $|\Phi(0)|^2=1$，从而必定有 $\Phi(0)=\pm 1$. 为了方便起见，我们选择 $\Phi(0)=1$. 由此得到规范化

$$\int_{-\infty}^{\infty}\phi(t)\mathrm{d}t=\Phi(0)=1 \tag{468b}$$

(Daubechies, 1992, p. 175). 所以，式(466a)可简化为

$$\Phi(f)=\prod_{m=1}^{\infty}\frac{\bar{G}\left(\frac{f}{2^m}\right)}{\sqrt{2}}. \tag{468c}$$

两尺度关系和 $\phi_{0,k}(\cdot)$的规范正交性同样可以用来证明滤波器$\{\bar{g}_l\}$具有单位能量以及它的偶数平移的正交性. 由以下练习指出(参看练习[76b]).

练习[468b] 证明

$$\sum_{l=-\infty}^{\infty}\bar{g}_l\,\bar{g}_{l+2k}=\begin{cases}1, & k=0;\\0, & \text{其他}.\end{cases} \tag{468d}$$

◁

11.5 尺度函数与尺度滤波器

本节，我们阐述感兴趣的尺度函数 $\phi(\cdot)$ 与第 j 层相应尺度滤波器 $\{g_{j,l}\}$（该滤波器在 4.6 节中已讨论，其一些主要的性质由表 154 给出）之间的关系. 因为 $g_l \equiv \bar{g}_{-l}$，所以第一层的尺度滤波器 $\{g_l\}$ 有如下的传递函数：

$$G(f) \equiv \sum_{l=-\infty}^{\infty} g_l e^{-i2\pi fl} = \sum_{l=-\infty}^{\infty} \bar{g}_{-l} e^{-i2\pi fl} = \bar{G}(-f).$$

在 5.2 节中，我们定义了极大重叠离散小波变换的尺度滤波器 $\tilde{g}_l \equiv g_l/\sqrt{2}$，它的传递函数是 $\tilde{G}(f) \equiv G(f)/\sqrt{2}$. 借助于这个传递函数，我们可以将式(468c)写为

$$\Phi(f) = \prod_{m=1}^{\infty} \tilde{G}\left(-\frac{f}{2^m}\right).$$

这一结论能使我们在 $\phi(\cdot)$ 与 $\{g_l\}$ 之间建立一种关系. 我们首先注意

$$\bar{\Phi}(f) \equiv \Phi(-f) = \prod_{m=1}^{\infty} \tilde{G}\left(\frac{f}{2^m}\right) \approx \prod_{m=1}^{j} \tilde{G}\left(\frac{f}{2^m}\right),$$

对于 j 很大时成立，其中这个逼近的程度可以用诸如 Daubechies 尺度滤波器（直观地，因为 $\tilde{G}(0)=1$ 和 $\tilde{G}(\cdot)$ 是一个对这些滤波器具有所有阶导数的连续函数，所以以上是合理的）来说明. 因此，由式(169b)有

$$\bar{\Phi}(2^j f) \approx \prod_{m=1}^{j} \tilde{G}(2^{j-m} f) = \prod_{l=0}^{j-1} \tilde{G}(2^l f) = \tilde{G}_j(f),$$

其中 $\tilde{G}_j(f)$ 为 $\{\tilde{g}_{j,l}\}$ 的传递函数. $\tilde{G}_j(f)$ 的离散傅里叶逆变换（式(35a)是 $\{\tilde{g}_{j,l}\}$，所以如果对上式两边进行离散傅里叶逆变换，并且考虑到 j 比较大时的情况，我们得到

$$\begin{aligned} \tilde{g}_{j,l} &\approx \int_{-1/2}^{1/2} \bar{\Phi}(2^j f) e^{i2\pi fl} \, df = \frac{1}{2^j} \int_{-2^{j-1}}^{2^{j-1}} \bar{\Phi}(f') e^{i2\pi f'(l/2^j)} \, df' \\ &\approx \frac{1}{2^j} \int_{-\infty}^{\infty} \bar{\Phi}(f') e^{i2\pi f'(l/2^j)} \, df' = \frac{1}{2^j} \phi\left(-\frac{l}{2^j}\right). \end{aligned}$$

因此我们得到结论

$$2^j \tilde{g}_{j,-l} = 2^{j/2} g_{j,-l} \approx \phi\left(\frac{l}{2^j}\right), \tag{469}$$

所以，对于大的 j，第 j 层反时间的等价的尺度滤波器与 $\phi(\cdot)$ 有相同的形状.

为了验证式(469)中的逼近，我们需要在网格 $l/2^j$ 上计算 $\phi(\cdot)$. 事实上，我们可以使用两尺度方程来计算尺度函数（在某一比例常数内通过在网格上递

推). 为此，让我们用 D(4)尺度滤波器为例来说明，我们有(通过式(75d))

$$\bar{g}_0 = g_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, \quad \bar{g}_{-1} = g_1 = \frac{3+\sqrt{3}}{4\sqrt{2}},$$

$$\bar{g}_{-2} = g_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, \quad \bar{g}_{-3} = g_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}.$$

当 $L=4$ 时，$\phi(\cdot)$的支撑为$(-(L-1),\ 0]=(-3,\ 0]$. 由两尺度差分方程(式(465a))，我们得到

$$\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\\ \phi(0)\end{bmatrix} = \sqrt{2}\begin{bmatrix}\bar{g}_{-3} & 0 & 0 & 0\\ \bar{g}_{-1} & \bar{g}_{-2} & \bar{g}_{-3} & 0\\ 0 & \bar{g}_0 & \bar{g}_{-1} & \bar{g}_{-2}\\ 0 & 0 & 0 & \bar{g}_0\end{bmatrix}\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\\ \phi(0)\end{bmatrix}.$$

从下面一行，我们有 $\phi(0)=\sqrt{2}\bar{g}_0\phi(0)$，因为$\sqrt{2}\bar{g}_0\neq 1$，这告诉我们 $\phi(0)=0$. 从而可以删除最后一行和列，并利用$\bar{g}_{-l}=g_l$，写成

$$\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix} = \sqrt{2}\begin{bmatrix}g_3 & 0 & 0\\ g_1 & g_2 & g_3\\ 0 & g_0 & g_1\end{bmatrix}\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix} = \boldsymbol{G}_0\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix}. \tag{470a}$$

类似地，我们也有 $\phi(-3)=0$(这与我们假设 $t=-3$ 在 $\phi(\cdot)$的支撑区间以外相一致). 从而我们可以删除第一行和列，得到

$$\begin{bmatrix}\phi(-2)\\ \phi(-1)\end{bmatrix} = \sqrt{2}\begin{bmatrix}g_2 & g_3\\ g_0 & g_1\end{bmatrix}\begin{bmatrix}\phi(-2)\\ \phi(-1)\end{bmatrix}.$$

以上说明向量$[\phi(-2),\ \phi(-1)]^T$ 是一 2×2 矩阵的特征向量(相应特征值为 1). 因为对于任一常数 c，$c[\phi(-2),\ \phi(-1)]^T$ 也是一特征向量，所有我们仅得到在某一比例常数下的 $\phi(-2)$和 $\phi(-1)$(实际应用中，利用 $\phi(\cdot)$必须是积分为 1 这一事实，以后我们可以非常精确地计算这一常数). 上面第一行给出了

$$\frac{\phi(-2)}{\phi(-1)} = \frac{g_3\sqrt{2}}{1-g_2\sqrt{2}} = \frac{1-\sqrt{3}}{1+\sqrt{3}}, \tag{470b}$$

因为 $\phi(-3)=0$，从而也就固定了$[\phi(-3),\ \phi(-2),\ \phi(-1)]^T$ 的相关元素大小. 假设现在我们在每个整数 -3，-2，-1 上都增加 0.5(二进制中表示为 .1)来估计 $\phi(\cdot)$在这些点上的值. 如果我们再一次使用两尺度方程，就得到

$$\begin{bmatrix}\phi(-2.5)\\ \phi(-1.5)\\ \phi(-0.5)\end{bmatrix} = \sqrt{2}\begin{bmatrix}g_2 & g_3 & 0\\ g_0 & g_1 & g_2\\ 0 & 0 & g_0\end{bmatrix}\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix} = \boldsymbol{G}_1\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix},$$

即由在整数点的值，得到两整数之间的值. 现在假设我们在每个整数元素 -3，-2，-1 上都增加 0.25(二进制中表示为 0.01)来估计 $\phi(\cdot)$在这些点上的值. 两尺度方程给出

$$\begin{bmatrix}\phi(-2.75)\\ \phi(-1.75)\\ \phi(-0.75)\end{bmatrix} = \boldsymbol{G}_0\begin{bmatrix}\phi(-2.5)\\ \phi(-1.5)\\ \phi(-0.5)\end{bmatrix} = \boldsymbol{G}_0\boldsymbol{G}_1\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix};$$

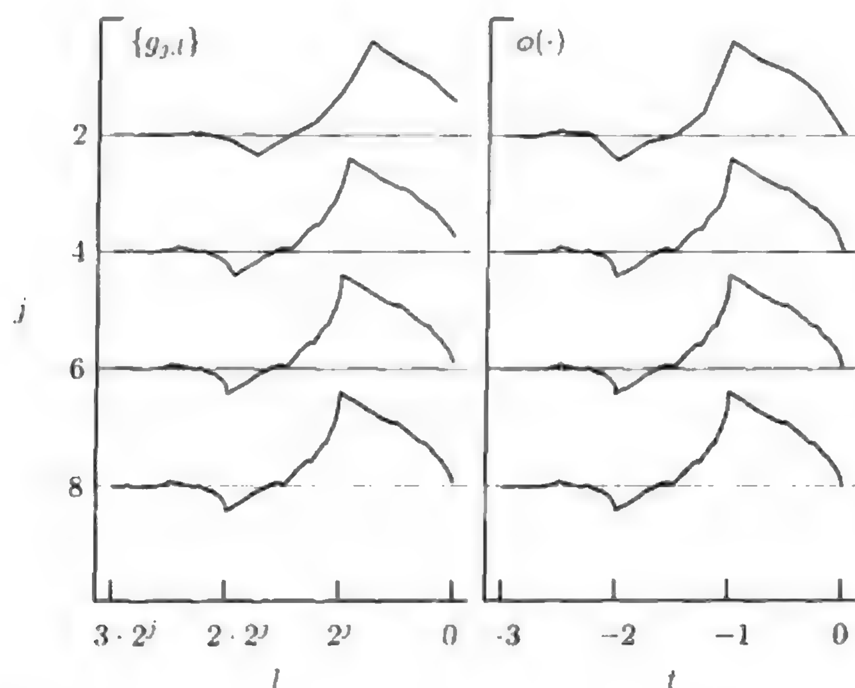


图 471　第 j 层等效的 D(4)尺度滤波器 $\{g_{j,l}\}$（左边一列）和 D(4)尺度函数$\phi(\cdot)$在网格$\frac{l}{2^j}$，$l=-3\cdot 2^j$，…，-1，0（右边一列），$j=2$，4，6，8（从上到下）上的估计值．对给定的 j，两组图像均由 $3\cdot 2^j+1$ 段直线段连接而成．右边一列从上到下分别表示的是函数 $\phi(\cdot)$在 $t=-3$ 到 $t=0$ 之间以步长为 0.25，0.062 5，0.015 625，0.003 906 25 的图像．左边一列中，滤波器表示了式(469)的近似，它的精确度随着 j 的增大而增高

类似地，如果在每个整数元素上增加 0.75(二进制中表示为 0.11)．两尺度方程现在给出

$$\begin{bmatrix}\phi(-2.25)\\ \phi(-1.25)\\ \phi(-0.25)\end{bmatrix}=\boldsymbol{G}_1\begin{bmatrix}\phi(-2.5)\\ \phi(-1.5)\\ \phi(-0.5)\end{bmatrix}=\boldsymbol{G}_1\boldsymbol{G}_1\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix}.$$

注意数量的二进制表示(添加整数位)告诉我们如何使用矩阵 $\boldsymbol{G}_0$ 和 $\boldsymbol{G}_1$ 的下标序列．例如，如果我们增加 0.625(二进制中表示为 0.101)，就使用 $\boldsymbol{G}_1\boldsymbol{G}_0\boldsymbol{G}_1$ 左乘向量$[\phi(-3),\ \phi(-2),\ \phi(-1)]^{\mathrm{T}}$来得到$[\phi(-2.375),\ \phi(-1.375),\ \phi(-0.375)]^{\mathrm{T}}$(更多递推格式的讨论参见 Strang and Nguyen，1996，p. 196)．

作为例子，图 471 右列从上到下显示了序列$\left\{\phi\left(\frac{l}{2^j}\right):l=-3\cdot 2^j,\ \cdots,\ -1,\ 0\right\}$中的 $j=2$，4，6，8．最上面的曲线是对 $\phi(\cdot)$的以 0.25 为步长的从 -3 到 0 的估计值，最下面的曲线是步长为 0.003 906 25．随着步长的减小，图像曲线慢慢的汇聚于 D(4)尺度函数特有的“鲨鱼鳍”形状．左边一列显示了与 $l=3\cdot 2^j$，…，1，0 相应的 $g_{j,l}$的图像(对于 $j=2$，4，6，8 在图 98a 中的左边一列也以相反的顺序同样给出了)．注意在 $j=2$ 和 4 时，左边曲线与右边曲线存在着视觉上的差别，在 $j=6$ 和 8 时这种差别更大，这正好验证了式(469)的逼近程度．

11.5 节的评论与扩展

[1]因为$\{g_{j,l}\}$具有宽度$L_j=(2^j-1)(L-1)+1$，我们知道$g_{j,l}=0$，对于$l<0$和$l>(2^j-1)(L-1)$成立．当j很大时，由式(469)有$2^{j/2}g_{j,l}\approx\phi\left(-\frac{l}{2^j}\right)$；然而$-\frac{1}{2^j}(2^j-1)(L-1)=-(L-1)+\frac{L-1}{2^j}$，对于很大的$j$，这刚好大于$-(L-1)$．由此论证能够得出结论$\phi(\cdot)$在$(-(L-1),0]$之外为0．式(460c)定义的哈尔尺度函数有$\phi(-1)=\phi(-(L-1))=0$和$\phi(0)=1$；另一方面，对D(4)尺度函数，使用两尺度关系，有$\phi(-3)=\phi(-(L-1))=0$而$\phi(0)=0$．因此，一个有用的通用表达为$\text{support}\{\phi(\cdot)\}\subset(-(L-1),0]$，这允许我们处理$\phi(0)=1$与$\phi(0)=0$的可能性(比较 Daubechies，1992，第6章)．

11.6 小波函数和细节空间

假设V_j，$j\in\mathbf{Z}$形成一个多分辨分析(在11.2节最后已定义)，所以特别地，$V_0\subset V_{-1}$．令$W_0\subset V_{-1}$是V_0在V_{-1}中的正交补；即如果$\varphi(\cdot)\in W_0$，那么$\varphi(\cdot)\in V_{-1}$但$\varphi(\cdot)$与所有函数$x(\cdot)\in V_0$正交，

$$\langle\varphi(\cdot),x(\cdot)\rangle=0.$$

然后可以定义

$$V_{-1}=V_0\oplus W_0,$$

即空间V_{-1}称为V_0和W_0的**直接和**，其意思是V_{-1}中的任一元素能够表示为两个相互正交元素之和，一个属于V_0，另一个属于W_0．一般地，$V_j\subset V_{j-1}$，且$W_j\subset V_{j-1}$是V_j在V_{j-1}中的正交补，所以有$V_j=V_{j-1}\oplus W_{j-1}$，由以上迭代我们得到

$$V_j=W_{j+1}\oplus W_{j+2}\oplus\cdots.$$

子空间W_j称为对尺度$\tau_j=2^{j-1}$的细节空间．V_j和W_j的嵌套如图473a所示．

作为例子，在图473b中，从左到右依次显示的是基于哈尔尺度函数在子空间$V_0^{(H)}$，$W_0^{(H)}$和$V_{-1}^{(H)}$中的一个函数．回想一下，所有$V_0^{(H)}$中的函数在区间$k-1<t\leq k$，$k\in\mathbf{Z}$上都为常数；另一方面，所有$V_{-1}^{(H)}$中的函数在区间$\frac{k-1}{2}<t\leq\frac{k}{2}$，$k\in\mathbf{Z}$上都为常数．要使函数$\varphi(\cdot)$在$V_{-1}^{(H)}$同时与$V_0^{(H)}$中的所有函数正交，在区间$\frac{k-2}{2}<t\leq\frac{k-1}{2}$和$\frac{k-1}{2}<t\leq\frac{k}{2}$($k$为偶数)上这两组值之和为0．这在图473b的中间图像中已表明，其中$\varphi(\cdot)$的值表示偏离0的程度，区间$k-1<t\leq k$用竖直的点线来给出．事实上，如果将图473b中左边和中间的图像求和，则得到右边的图像．

让我们定义小波函数$\psi(\cdot)$为

$$\psi(t)\equiv\sqrt{2}\sum_{l=-\infty}^{\infty}\bar{h}_l\phi(2t-l),\quad \text{其中}\ \bar{h}_l\equiv(-1)^l\bar{g}_{1-l-L}. \tag{472}$$

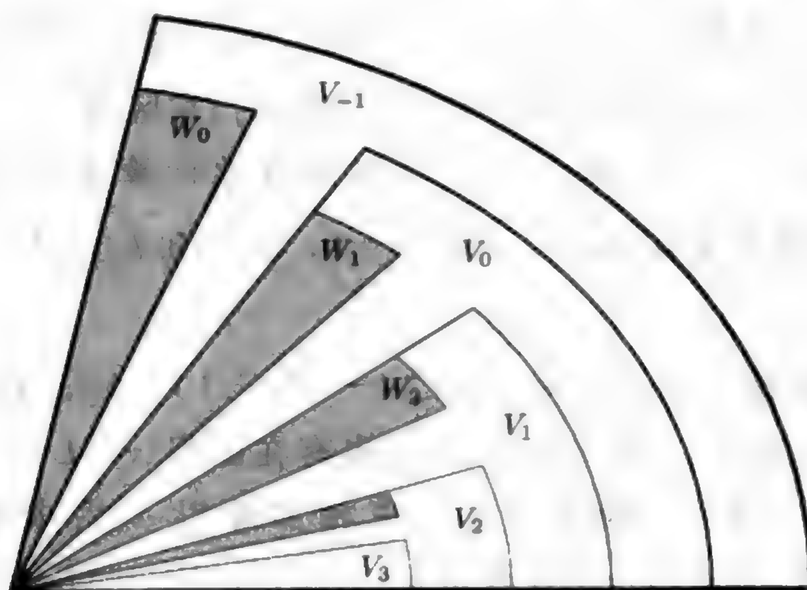


图 473a 韦恩图阐明了子空间 V_j 和 W_j 的嵌套. 有五段圆弧从基线发出. 每段圆弧的末端有一段直线，连接到左边较低角点. 由给定弧和从弧的末端发射的两条直线所围成的面积表示一个近似空间 V_j. 这最大面积的整个图表示 V_{-1}，而最小面积的图像表示 V_3(注意按要求 $V_3 \subset V_2 \subset V_1 \subset V_0 \subset V_{-1}$). 阴影部分表示细节空间 W_j. 注意 $W_0 \subset V_{-1}$，$W_1 \subset V_0$，$W_2 \subset V_1$，$W_3 \subset V_2$(因为写不下，所以 W_3 没有标出). 同时注意 $V_0 \subset V_{-1}$ 和 $W_0 \subset V_{-1}$，因为 V_{-1} 同时还包含 W_0 和 V_0 之内的函数的线性组合——这种组合不一定要同时在 V_0 和 W_0 中，可以在 V_0 和 W_0 之外的空间中，在图中以镰刀状的图形表示，并标以 V_{-1}，所以 $V_0 \cup W_0 \neq V_{-1}$. 最后注意因为所有的空间都包括零函数，所以所有的 V_j 和 W_j 都相交于一点(左下角的交点)

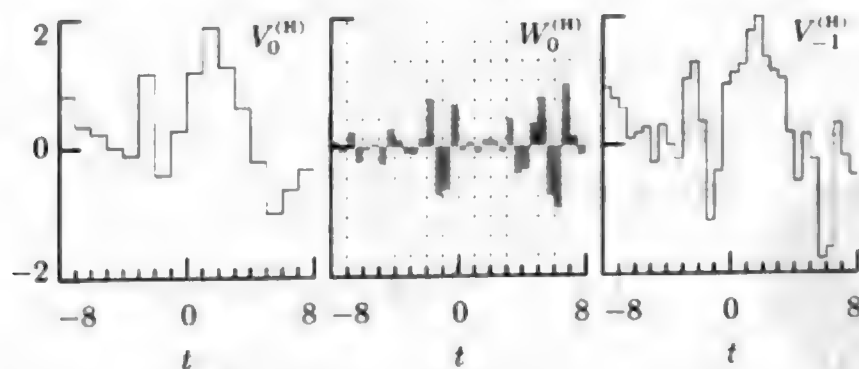


图 473b 在空间 $V_0^{(H)}$，$W_0^{(H)}$ 和 $V_{-1}^{(H)}$ 中的函数举例. 注意当 $V_0^{(H)}$ 空间中的函数在区间$(k-1, k]$，$k \in \mathbf{Z}$，中为常数时，则在这一区间中，$W_0^{(H)}$ 中的函数的积分为零(这些区间的端点在中间图像显示垂直点线显示). 在 $V_{-1}^{(H)}$ 中的函数事实上由另外两个图像中由点法则相加形成

(注意这一定义与式(465a)中关于$\phi(\cdot)$的两尺度关系之间的相似性). 序列$\{\bar{h}_l\}$是相应于$\{\bar{g}_l\}$的正交镜像滤波器(QMF); 正交镜像滤波器的定义隐含, 假设$\{\bar{g}_l\}$在区间$-(L-1)\leqslant l\leqslant 0$之外等于0, 因此$\{\bar{h}_l\}$在区间$-(L-1)\leqslant l\leqslant 0$之外也等于0. 因此, 从$-\infty$到$\infty$的求和在数学计算方面很方便解释且与式(465a)一致, 但是在实际应用中, 取$\bar{h}_l=0$, $l>0$和$l\leqslant -L$. 正交镜像滤波器的定义保证$\bar{h}_{-l}=h_l$.

考虑到支撑区间 support$\{\phi(\cdot)\}\subset(-(L-1), 0]$, 式(472)意味着当$-(L-1)<2t-l\leqslant 0$时, $\psi(t)=0$. 我们也有在区间$-(L-1)\leqslant l\leqslant 0$之外有$\{\bar{h}_l\equiv(-1)^l\bar{g}_{1-l-L}\}$为0. 因此, 两约束相加除以2后求得, 当$t$不属于$-(L-1)<t\leqslant 0$时$\psi(t)=0$, 所以有 support$\{\psi(\cdot)\}\subset(-(L-1), 0]$, 对$\phi(\cdot)$也一样.

带有

$$\psi_{j,k}(t)\equiv\frac{\psi_{0,k}\left(\frac{t}{2^j}\right)}{\sqrt{2^j}}=\frac{\psi\left(\frac{t}{2^j}-k\right)}{\sqrt{2^j}},\quad j,k\in\mathbf{Z},$$

我们要求$\{\psi_{0,m}(\cdot): m\in\mathbf{Z}\}$构成细节空间$W_0$的一个规范正交基. 为此, 我们先证明, 对所有的m, $n\in\mathbf{Z}$有$\psi_{0,m}(\cdot)\perp\phi_{0,n}(\cdot)$, 其中$\perp$表示"正交于", 即

$$\langle\psi_{0,m}(\cdot),\phi_{0,n}(\cdot)\rangle\equiv\int_{-\infty}^{\infty}\psi_{0,m}(t)\phi_{0,n}(t)\,\mathrm{d}t=0.$$

这一结论可从$\phi_{0,n}$的规范正交性和式(465a)以及(472)得出, 因为我们有

$$\int_{-\infty}^{\infty}\psi_{0,m}(t)\phi_{0,n}(t)\,\mathrm{d}t=\sum_{l=-\infty}^{\infty}\bar{h}_l\,\bar{g}_{l+2m-2n}=0.$$

练习[474a] 补充上面等式的推导细节. ◁

因为$\phi_{0,n}(\cdot)$构成V_0的一个基底, 由此得出$\psi_{0,m}(\cdot)$与V_0中的所有函数正交. 另外, 式(472)中$\psi(\cdot)$的定义可以写成

$$\psi_{0,m}(t)=\sum_{l=-\infty}^{\infty}\bar{h}_l\phi_{-1,l}(t-m),$$

这表明$\psi_{0,m}(\cdot)\in V_{-1}$. 由于$\psi_{0,m}(\cdot)$与$V_0$中的所有函数正交, 且属于$V_{-1}$, 所以我们必有$\psi_{0,m}(\cdot)\in W_0$对所有的$m\in\mathbf{Z}$成立.

练习[474b] 使用$\psi(\cdot)$的定义, $\phi_{0,m}(\cdot)$的规范正交性和由式(468d)给出的$\{g_l\}$的偶数平移的正交性, 证明$\psi_{0,m}(\cdot)$是规范正交的. ◁

最后一步需要建立$\psi_{0,m}(\cdot)$构成W_0的一个规范正交基, 这是练习[11.2]的主题. 因此, 像V_j的基函数是尺度函数$\phi(\cdot)$的伸缩平移一样, W_j的基函数是小波函数$\psi(\cdot)$的伸缩平移.

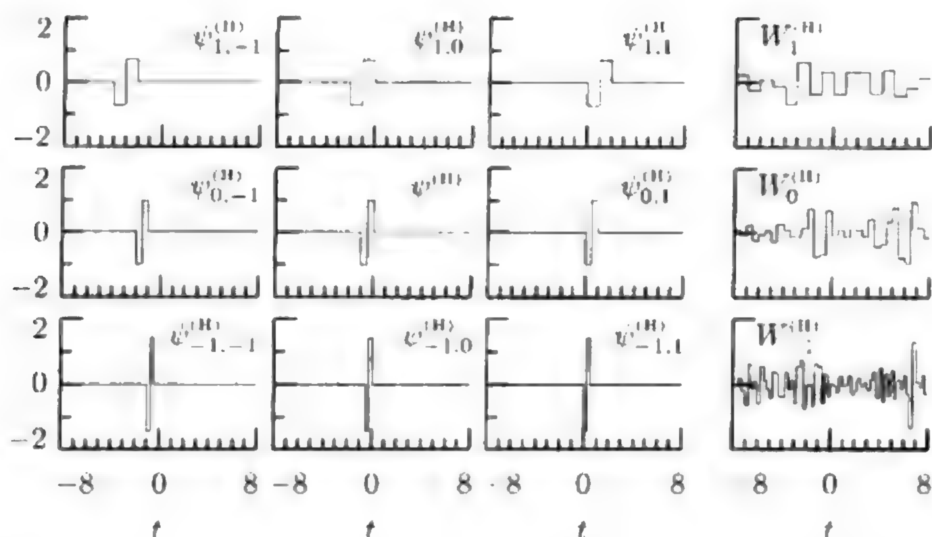


图 475　哈尔小波函数 $\psi^{(H)}(\cdot)$与相应的细节空间. 中间一行的前三个显示的是哈尔细节空间 $W_0^{(H)}$的三个基函数，即从左到右，$\psi_{0,-1}^{(H)}(\cdot)$，$\psi^{(H)}(\cdot)$和$\psi_{0,1}^{(H)}(\cdot)$. 这行最右边的图像是 $W_0^{(H)}$中的一个函数的例子. 上排和下排分别显示了与哈尔细节空间 $W_1^{(H)}$和 $W_{-1}^{(H)}$相应的图像(图 461 显示了相应的哈尔尺度函数和逼近空间)

小波函数的一个最简单的例子是哈尔小波函数，其定义如下：

$$\psi^{(H)}(t) \equiv \sqrt{2}\sum_{l=-\infty}^{\infty}\bar{h}_l\phi^{(H)}(2t-l) = \sqrt{2}\left[-\frac{\phi^{(H)}(2t+1)}{\sqrt{2}} + \frac{\phi^{(H)}(2t)}{\sqrt{2}}\right]$$

$$= \begin{cases} -1, & -1 < t \leqslant -1/2; \\ 1, & -1/2 < t \leqslant 0; \\ 0, & \text{其他}. \end{cases} \tag{475}$$

由于正交镜像滤波器关系产生 $\bar{h}_{-1} = -\bar{g}_0 = -1/\sqrt{2}$，$\bar{h}_0 = \bar{g}_{-1} = 1/\sqrt{2}$ 和 $\bar{h}_l = 0$($l \neq 0$ 或-1)(注意这一定义与式(2c)中定义的不同). 图 475 显示了哈尔小波基函数，从上到下依次对于子空间 $W_1^{(H)}$，$W_0^{(H)}$ 和 $W_{-1}^{(H)}$的，以及每一个子空间的一个函数实例(比较图 461 的哈尔尺度函数).

因为 $\psi_{1,k} \in W_1$，$\phi_{1,m}(\cdot) \in V_1$，且由于 W_1中所有函数都与 V_1中的函数正交，我们有 $\psi_{1,k}(\cdot) \perp \phi_{1,m}(\cdot)$，对所有 k，$m \in \mathbf{Z}$ 成立，进而因为

$$V_0 = V_2 \oplus W_2 \oplus W_1,$$

则有 $\phi_{2,k}(\cdot) \perp \psi_{2,m}(\cdot)$，$\phi_{2,k}(\cdot) \perp \psi_{1,m}(\cdot)$以及 $\psi_{2,k}(\cdot) \perp \psi_{1,m}(\cdot)$，对所有 k，$m \in \mathbf{Z}$ 成立. 一般地，$\phi_{j,k}(\cdot) \perp \psi_{l,m}(\cdot)$，$j$，$k$，$l$，$m \in \mathbf{Z}$ 且使 $l \leqslant j$. $\psi_{j,k}(\cdot) \perp \psi_{l,m}(\cdot)$，对于所有 j，k，l，$m \in \mathbf{Z}$ 使 $j \neq l$ 以及 $\psi_{j,k}(\cdot) \perp \psi_{j,m}(\cdot)$，对于所有 j，k，$m \in \mathbf{Z}$ 使 $k \neq m$.

$\{\bar{g}_l\}$和$\{\bar{h}_l\}$之间的正交镜像滤波器关系可以用于建立如下结果.

练习[476]　证明$\{\bar{h}_l\}$的傅里叶变换由下式给出：

$$\bar{H}(f) = -\,\mathrm{e}^{\mathrm{i}2\pi f(L-1)}\bar{G}^*\left(f-\frac{1}{2}\right) = -\,\mathrm{e}^{\mathrm{i}2\pi f(L-1)}\bar{G}\left(\frac{1}{2}-f\right). \tag{476a}$$

◁

由式(468a)和(476a)以及$\bar{G}(\cdot)$的单位周期性产生

$$\bar{H}(0) = \sum_{l=-\infty}^{\infty}\bar{h}_l = 0. \tag{476b}$$

另外，令 $\Psi(\cdot)$表示 $\psi(\cdot)$的傅里叶变换. 用类似式(465a)到(465b)导出的论证，对式(472)两边进行傅里叶变换得到

$$\Psi(f) = \Phi\left(\frac{f}{2}\right)\frac{\bar{H}\left(\frac{f}{2}\right)}{\sqrt{2}}. \tag{476c}$$

因此 $\bar{H}(0)=0$ 意味着

$$\Psi(0) = \int_{-\infty}^{\infty}\psi(t)\,\mathrm{d}t = 0.$$

(Daubechies，1992，2.4 节). 简单地讲，以上的意思就是说小波函数看起来像一个"小的波"；因为 $\Psi(0)=0$，小波函数在频率为零的时候没有能量. 因此，像尺度函数一样，小波函数 $\psi(\cdot)$具有单位范数，但是与尺度函数不同的是，它的积分为零. 将式(465b)代入式(476c)，回忆 $\Phi(0)=1$ 我们有

$$\Psi(f) = \frac{\bar{H}\left(\frac{f}{2}\right)}{\sqrt{2}}\prod_{m=2}^{\infty}\frac{\bar{G}\left(\frac{f}{2^m}\right)}{\sqrt{2}}. \tag{476d}$$

11.7　小波函数与小波滤波器

在小波函数 $\psi(\cdot)$与第 j 层等价小波滤波器$\{h_{j,l}\}$之间，存在类似于我们在 11.5 节中讲的尺度函数与滤波器之间的关系. 因为 $h_l \equiv \bar{h}_{-l}$，我们知道 $H(f)=\bar{H}(-f)$. 因为极大重叠离散小波变换小波滤波器是$\tilde{h}_l \equiv h_l/\sqrt{2}$(参见 5.2 节)，它的传递函数由 $\tilde{H}(f)\equiv H(f)/\sqrt{2}$给出. 因此，式(476d)给出

$$\Psi(f) = \tilde{H}\left(-\frac{f}{2}\right)\prod_{m=2}^{\infty}\tilde{G}\left(-\frac{f}{2^m}\right),$$

从而，对于大的 j，

$$\bar{\Psi}(f) \equiv \Psi(-f) = \tilde{H}\left(\frac{f}{2}\right)\prod_{m=2}^{\infty}\tilde{G}\left(\frac{f}{2^m}\right) \approx H\left(\frac{f}{2}\right)\prod_{m=2}^{j}\tilde{G}\left(\frac{f}{2^m}\right).$$

因此，由式(169b)

$$\bar{\Psi}(2^j f) \approx \tilde{H}(2^{j-1}f)\prod_{m=2}^{j}\tilde{G}(2^{j-m}f) = \tilde{H}(2^{j-1}f)\prod_{l=0}^{j-2}\tilde{G}(2^l f) = \tilde{H}_j(f),$$

其中 $\widetilde{H}_j(\cdot)$ 是 $\{\widetilde{h}_{j,l}\}$ 的传递函数. 对于大的 j，由离散傅里叶逆变换产生

$$\widetilde{h}_{j,l} \approx \int_{-1/2}^{1/2} \overline{\Psi}(2^j f) e^{i2\pi fl} df \approx \frac{1}{2^j}\psi\left(-\frac{l}{2^j}\right).$$

所以对于大的 j，

$$2^j \widetilde{h}_{j,-l} = 2^{j/2} h_{j,-l} \approx \psi\left(\frac{l}{2^j}\right), \tag{477a}$$

因此第 j 层等价小波滤波器的时间逆将与 $\psi(\cdot)$ 有近似的图形.

现在，让我们转到利用式(472)的递推来计算 $\psi(\cdot)$. 举一个特殊的例子，利用式(59a)定义的小波滤波器来再一次考虑 D(4)的设计. 由式(472)以及用 $h_l = \overline{h}_{-l}$ 得到

$$\begin{bmatrix}\psi(-3)\\ \psi(-2)\\ \psi(-1)\\ \psi(0)\end{bmatrix} = \sqrt{2}\begin{bmatrix} h_3 & 0 & 0 & 0\\ h_1 & h_2 & h_3 & 0\\ 0 & h_0 & h_1 & h_2\\ 0 & 0 & 0 & h_0\end{bmatrix}\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\\ \phi(0)\end{bmatrix}.$$

因为我们知道 $\phi(-3)=\phi(0)=0$，由此得出 $\psi(-3)=\psi(0)=0$. 删除第一行，第一列和最后一行最后一列得到

$$\begin{bmatrix}\psi(-2)\\ \psi(-1)\end{bmatrix} = \sqrt{2}\begin{bmatrix} h_2 & h_3\\ h_0 & h_1\end{bmatrix}\begin{bmatrix}\phi(-2)\\ \phi(-1)\end{bmatrix}, \tag{477b}$$

由此我们可以得出比率

$$\frac{\psi(-2)}{\psi(-1)} = -\frac{1+\sqrt{3}}{1-\sqrt{3}} \tag{477c}$$

及 $[\psi(-3), \psi(-2), \psi(-1)]^T$ 的元素的相对值(比率的验证是练习[11.3]的主题). 类似于式(470a)，我们可以写

$$\begin{bmatrix}\psi(-3)\\ \psi(-2)\\ \psi(-1)\end{bmatrix} = \sqrt{2}\begin{bmatrix} h_3 & 0 & 0\\ h_1 & h_2 & h_3\\ 0 & h_0 & h_1\end{bmatrix}\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix} = \boldsymbol{H}_0\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix}.$$

现在假设对整数 -3，-2，-1 分别增加 0.5(二进制表示为 .01)，在这些点上估计 $\psi(\cdot)$ 的值. 再一次利用式(472)得

$$\begin{bmatrix}\psi(-2.5)\\ \psi(-1.5)\\ \psi(-0.5)\end{bmatrix} = \sqrt{2}\begin{bmatrix} h_2 & h_3 & 0\\ h_0 & h_1 & h_2\\ 0 & 0 & h_0\end{bmatrix}\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix} = \boldsymbol{H}_1\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix}.$$

如果对整数 -3，-2，-1 分别增加 0.25(二进制表示为 0.01)，在这些点上估计 $\psi(\cdot)$ 的值. 式(472)给出

$$\begin{bmatrix}\psi(-2.75)\\ \psi(-1.75)\\ \psi(-0.75)\end{bmatrix} = \boldsymbol{H}_0\begin{bmatrix}\phi(-2.5)\\ \phi(-1.5)\\ \phi(-0.5)\end{bmatrix} = \boldsymbol{H}_0\boldsymbol{G}_1\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix}.$$

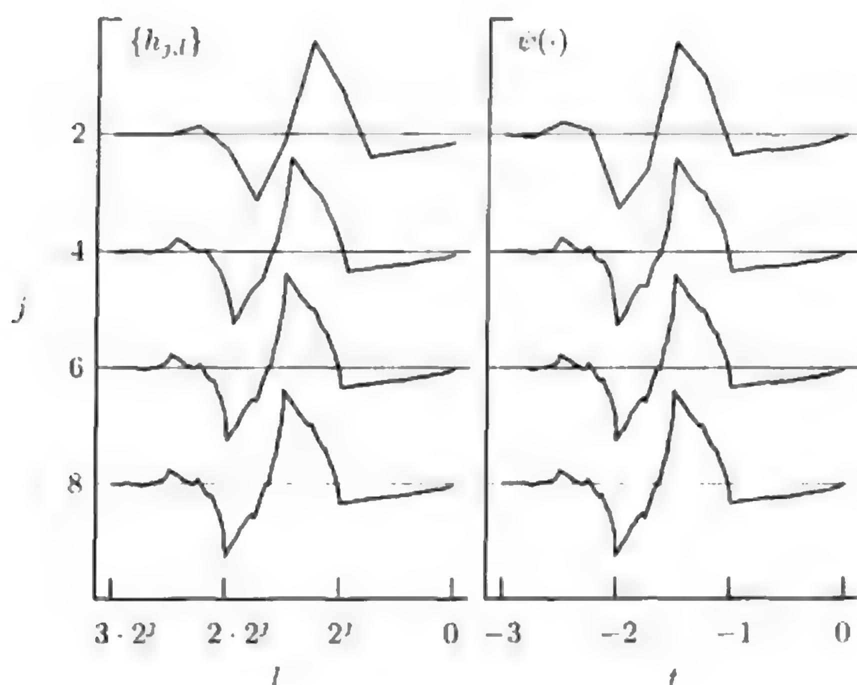


图 478 第 j 阶等价的 D(4)小波滤波器 $\{h_{j,l}\}$（左边一列）和 D(4)小波函数 $\psi(\cdot)$ 在网格 $\frac{l}{2^j}$，$l=-3\cdot 2^j$，…，-1，0（右边一列），$j=2$，4，6，8（从上到下）上的估计值．详细的布局参见图 471 的类似的说明

类似地，如果对整数 −3，−2，−1 分别增加 0.75(二进制表示为 0.11)，在这些点上估计$\psi(\cdot)$的值．我们得到

$$\begin{bmatrix}\psi(-2.25)\\ \psi(-1.25)\\ \psi(-0.25)\end{bmatrix} = \boldsymbol{H}_1\begin{bmatrix}\phi(-2.5)\\ \phi(-1.5)\\ \phi(-0.5)\end{bmatrix} = \boldsymbol{H}_1\boldsymbol{G}_1\begin{bmatrix}\phi(-3)\\ \phi(-2)\\ \phi(-1)\end{bmatrix}.$$

二进制数的整数位增加的表示形式再一次告诉我们使用矩阵序列．例如，如果增加 0.625（二进制表示为 0.101），则用 $\boldsymbol{H}_1\boldsymbol{G}_0\boldsymbol{G}_1$ 乘以向量$[\phi(-3),\ \phi(-2),\ \phi(-1)]^{\mathrm{T}}$．最后一个矩阵乘法是 $\boldsymbol{H}_n$，所以之前将插入矩阵 $\boldsymbol{G}_n$.

作为例子，图 478 右边一列显示序列$\left\{\psi\left(\frac{l}{2^j}\right):l=-3\cdot 2^j,\ \cdots,\ -1,\ 0\right\}$(从上到下)对于 $j=2$，4，6，8，左边一列显示了与 $l=3\cdot 2^j$，…，1，0 相应的 $h_{j,l}$ 的图像($j=2$，4，6 的曲线在图 98a 的右边一列逆序显示)．随着 j 的增大，同一行的图像之间的视觉差异逐渐消失，由此验证式(477a)的逼近程度.

11.8 有限能量信号的多分辨分析

在 11.3 节中，我们考虑通过向子空间 $V_j\subset V_0$上投影获得的 $s_j(\cdot)$来逼近信号 $x(\cdot)\in V_0$．这里我们通过证明在 $x(\cdot)$与投影 $s_j(\cdot)$之间的差别能够表示为在细节子空间 W_j，W_{j-1}，…，W_1上的投影来展开我们的讨论．信号 $x(\cdot)\in$

$V_0=V_1\oplus W_1$能够分解为它在V_1和W_1上的分量给出和的形式(多分辨分析)：

$$x(t)=s_1(t)+d_1(t),$$

其中$s_1(\cdot)$和$d_1(\cdot)$分别为信号$x(\cdot)$包含在V_1和W_1上的部分．信号$s_1(\cdot)$是$x(\cdot)$在子空间V_1的投影，由式(462b)给出，同样地，信号$d_1(\cdot)$是$x(\cdot)$在子空间W_1的投影．因为$\{\psi_{1,k}(\cdot)\}$形成W_1的一个规范正交基，我们有

$$d_1(t)=\sum_{k=-\infty}^{\infty}w_{1,k}\psi_{1,k}(t),\text{其中 } w_{1,k}\equiv\langle x(\cdot),\psi_{1,k}(\cdot)\rangle. \tag{479a}$$

这里$d_1(\cdot)$表示$x(\cdot)$由粗糙的逼近$s_1(\cdot)$失掉的细节，$w_{1,k}$称为$d_1(\cdot)$的小波系数．使用式(463a)改进的平行结论，我们有

$$\begin{aligned}w_{1,k}&=\int_{-\infty}^{\infty}\Big[\sum_{m=-\infty}^{\infty}v_{0,m}\phi_{0,m}(t)\Big]\frac{\psi\left(\frac{t}{2}-k\right)}{\sqrt{2}}\mathrm{d}t\\&=\sum_{m=-\infty}^{\infty}v_{0,m}\int_{-\infty}^{\infty}\phi(t-[m-2k])\frac{\psi\left(\frac{t}{2}\right)}{\sqrt{2}}\mathrm{d}t\\&=\sum_{m=-\infty}^{\infty}v_{0,m}\langle\phi_{0,m-2k}(\cdot),\psi_{1,0}(\cdot)\rangle.\end{aligned} \tag{479b}$$

练习[479]　证明$\langle\phi_{0,m-2k}(\cdot),\ \psi_{1,0}(\cdot)\rangle=\bar{h}_{m-2k}$，其中$\{\bar{h}_l\}$是相应于$\{\bar{g}_l\}$的正交镜像滤波器．另外，证明

$$\langle\phi_{j,m}(\cdot),\psi_{j+1,k}(\cdot)\rangle=\langle\phi_{0,m-2k}(\cdot),\psi_{1,0}(\cdot)\rangle=\bar{h}_{m-2k} \tag{479c}$$

对于所有的j成立．◁

我们注意

$$\bar{h}_l=\langle\phi_{0,l}(\cdot),\psi_{1,0}(\cdot)\rangle=\int_{-\infty}^{\infty}\phi(t-l)\frac{\psi\left(\frac{t}{2}\right)}{\sqrt{2}}\mathrm{d}t. \tag{479d}$$

作为例子，让我们计算对于式(460c)和(475)的小波尺度$\phi^{(\mathrm{H})}(\cdot)$和小波函数$\psi^{(\mathrm{H})}(\cdot)$的$\{\bar{h}_l\}$．因为

$$\phi^{(\mathrm{H})}(t-l)=\begin{cases}1, & l-1<t\leqslant l;\\0, & \text{其他};\end{cases}\quad \psi^{(\mathrm{H})}\left(\frac{t}{2}\right)=\begin{cases}-1, & -2<t\leqslant -1;\\1, & -1<t\leqslant 0;\\0, & \text{其他}.\end{cases}$$

由式(479d)得到

$$\bar{h}_l=-\frac{1}{\sqrt{2}}\int_{-2}^{-1}\phi^{(\mathrm{H})}(t-l)\mathrm{d}t+\frac{1}{\sqrt{2}}\int_{-1}^{0}\phi^{(\mathrm{H})}(t-l)\mathrm{d}t=\begin{cases}-1/\sqrt{2}, & l=-1;\\1/\sqrt{2}, & l=0;\\0, & \text{其他}.\end{cases}$$

因为 $h_l \equiv h_{-l}$，我们得到

$$h_l = \begin{cases} 1/\sqrt{2}, & l = 0; \\ -1/\sqrt{2}, & l = 1; \\ 0, & \text{其他}. \end{cases}$$

这与4.2节中引入的哈尔小波滤波器 $\{h_l\}$ 的定义一致.

由式(479b)和练习[479]的第一部分，我们有

$$w_{1,k} = \sum_{m=-\infty}^{\infty} \bar{h}_{m-2k} v_{0,m} = \sum_{l=-\infty}^{\infty} h_l v_{0,2k-l}, \tag{480a}$$

其中我们使用 $h_l \equiv \bar{h}_{-l}$ 且令 $l = 2k - m$. 这个式与(463a)的 $v_{1,k}$ 非常类似. 像 $\{v_{1,k}\}$ 是正确的一样，在 $\{h_l\}$ 情况下，序列 $\{w_{1,k}\}$ 由对输出滤波器的(因子2)下采样产生.

因为 $s_1(\cdot) \in V_1 = V_2 \oplus W_2$，我们可以在 V_2 和 W_2 里将它分解得到

$$s_1(t) = s_2(t) + d_2(t), \quad d_2(t) \equiv \sum_{k=-\infty}^{\infty} w_{2,k} \psi_{2,k}(t),$$

其中 $j=2$ 层的小波系数由下式给出：

$$w_{2,k} \equiv \langle s_1(\cdot), \psi_{2,k}(\cdot) \rangle.$$

使用式(462b)中对 $s_1(\cdot)$ 的展式，我们有

$$\begin{aligned} w_{2,k} &= \langle \sum_{m=-\infty}^{\infty} v_{1,m} \phi_{1,m}(\cdot), \psi_{2,k}(\cdot) \rangle \\ &= \sum_{m=-\infty}^{\infty} v_{1,m} \langle \phi_{1,m}(\cdot), \psi_{2,k}(\cdot) \rangle \\ &= \sum_{m=-\infty}^{\infty} \bar{h}_{m-2k} v_{1,m} = \sum_{l=-\infty}^{\infty} h_l v_{1,2k-l}, \end{aligned} \tag{480b}$$

其中我们用到式(497c)当 $j=1$ 时的结果.

用类似的方法，我们可以将 $s_{j-1}(\cdot) \in V_{j-1} = V_j \oplus W_j$ 分解到 V_j 和 W_j 中的分量得到

$$s_{j-1}(t) = s_j(t) + d_j(t), \qquad d_j(t) \equiv \sum_{k=-\infty}^{\infty} w_{j,k} \psi_{j,k}(t), \tag{480c}$$

其中 j 层的小波系数由下式给出：

$$w_{j,k} \equiv \langle s_{j-1}(\cdot), \psi_{j,k}(\cdot) \rangle.$$

对 $w_{j,k}$ 的表达式进行递推得到

$$w_{j,k} = \sum_{m=-\infty}^{\infty} \bar{h}_{m-2k} v_{j-1,m} = \sum_{l=-\infty}^{\infty} h_l v_{j-1,2k-l} \tag{480d}$$

(参见式(464b)中的 $v_{j,k}$).

在第 j 步，我们可以写成

$$\begin{aligned} x(t) &= s_1(t) + d_1(t) \\ &= s_2(t) + d_2(t) + d_1(t) \\ &\vdots \\ &= s_j(t) + d_j(t) + \cdots + d_1(t), \end{aligned} \tag{481a}$$

一个多分辨分析反映这样一个事实

$$V_0 = V_j \oplus W_j \oplus \cdots \oplus W_1.$$

这个结果自然地导出重要的结果：$s_{j-1}(\cdot)$在 V_j 上的投影等于 $x(\cdot)$在 V_j 上的投影，$s_{j-1}(\cdot)$在 W_j 的投影等于 $x(\cdot)$在 W_j 上的投影. 下面的练习可以清楚地说明这个结论.

练习[481] 证明，$v_{2,k}$和 $w_{2,k}$的表达式由下式给出：

$$v_{2,k} = \langle x(\cdot), \phi_{2,k}(\cdot) \rangle, \qquad w_{2,k} = \langle x(\cdot), \psi_{2,k}(\cdot) \rangle.$$

推广这个结果，证明

$$v_{j,k} = \langle x(\cdot), \phi_{j,k}(\cdot) \rangle, \qquad w_{j,k} = \langle x(\cdot), \psi_{j,k}(\cdot) \rangle. \qquad \triangleleft$$

借助于尺度和小波系数，式(481a)中 $x(\cdot)$的表达式能够明确地写为

$$x(t) = \sum_{k=-\infty}^{\infty} v_{j,k}\phi_{j,k}(t) + \sum_{l=1}^{j}\sum_{k=-\infty}^{\infty} w_{l,k}\psi_{l,k}(t). \tag{481b}$$

在上式表达式中 $\phi_{j,k}(\cdot)$和 $\psi_{l,k}(\cdot)$的规范正交性容易得出帕塞瓦尔定理的如下形式：

$$\int_{-\infty}^{\infty} x^2(t)\mathrm{d}t = \sum_{k=-\infty}^{\infty} v_{j,k}^2 + \sum_{l=1}^{j}\sum_{k=-\infty}^{\infty} w_{l,k}^2. \tag{481c}$$

现在假设对某个 $j_0 \in \mathbf{Z}$，$x(\cdot) \in V_{j_0} = V_{j_0+1} \oplus W_{j_0+1}$；即我们没有必要像前面那样假设 $x(\cdot) \in V_0$，所以现在认为 $x(\cdot)$具有尺度 $\lambda_{j_0} = 2^{j_0}$ 而不是单位尺度. 一个类似上述的扩展显示，对任一的 $j > j_0$，我们有

$$x(t) = s_j(t) + d_j(t) + \cdots + d_{j_0+1}(t) = \sum_{k=-\infty}^{\infty} v_{j,k}\phi_{j,k}(t) + \sum_{l=j_0+1}^{j}\sum_{k=-\infty}^{\infty} w_{l,k}\psi_{l,k}(t), \tag{481d}$$

从而式(481c)变为

$$\int_{-\infty}^{\infty} x^2(t)\mathrm{d}t = \sum_{k=-\infty}^{\infty} v_{j,k}^2 + \sum_{l=j_0+1}^{j}\sum_{k=-\infty}^{\infty} w_{l,k}^2.$$

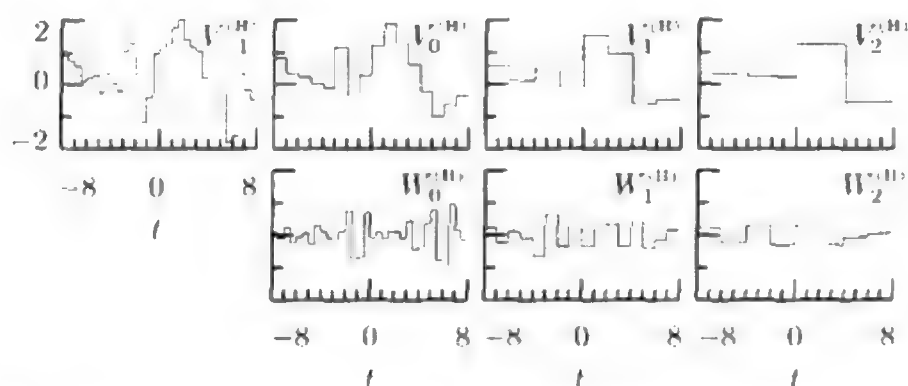


图 482 函数 $x(\cdot)\in V_{-1}^{(H)}$（左上角图）的多分辨分析，得到三个逼近，即 $s_0(\cdot)\in V_0^{(H)}$，$s_1(\cdot)\in V_1^{(H)}$ 和 $s_2(\cdot)\in V_2^{(H)}$（如图上排从左到右）以及相应的细节 $d_0(\cdot)\in W_0^{(H)}$，$d_1(\cdot)\in W_1^{(H)}$ 和 $d_2(\cdot)\in W_2^{(H)}$（下排从左到右）

作为例子，图 482 说明了"半尺度"函数 $x(\cdot)\in V_{-1}^{(H)}$ 的多分辨分析直到尺度 $\lambda_2=4$.

在式(481d)中，因为我们考虑 $x(\cdot)\in V_{j_0}\subset L^2(\mathbf{R})$，所以对 l 求和后是有限的. 我们仍然假设仅仅知道 $x(\cdot)\in L^2(\mathbf{R})$.

练习[482] 证明，对任一给定的 j，

$$\overline{V_j \oplus W_j \oplus W_{j-1}\cdots} = \overline{\bigcup_{l\in\mathbf{Z}} V_l} = L^2(\mathbf{R}). \qquad \triangleleft$$

由上式可得非齐次小波展开

$$x(t) = \sum_{k=-\infty}^{\infty} v_{j,k}\phi_{j,k}(t) + \sum_{l=-\infty}^{j}\sum_{k=-\infty}^{\infty} w_{l,k}\psi_{l,k}(t).$$

我们也可以证明

$$\overline{\bigoplus_{l\in\mathbf{Z}} W_l} = L^2(\mathbf{R}),$$

产生仅根据小波函数的齐次小波展开

$$x(t) = \sum_{l=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} w_{l,k}\psi_{l,k}(t).$$

最后，我们注意到，给定 j 层的尺度和小波系数，使用式(480c)，我们可以重构 $j-1$ 层的尺度系数

$$\begin{aligned} v_{j-1,k} &= \langle s_{j-1}(\cdot),\phi_{j-1,k}(\cdot)\rangle \\ &= \langle s_j(\cdot)+d_j(\cdot),\phi_{j-1,k}(\cdot)\rangle \end{aligned}$$

$$= \langle s_j(\cdot), \phi_{j-1,k}(\cdot)\rangle + \langle d_j(\cdot), \phi_{j-1,k}(\cdot)\rangle$$

$$= \sum_{m=-\infty}^{\infty} v_{j,m}\langle \phi_{j,m}(\cdot), \phi_{j-1,k}(\cdot)\rangle + \sum_{m=-\infty}^{\infty} w_{j,m}\langle \psi_{j,m}(\cdot), \phi_{j-1,k}(\cdot)\rangle$$

$$= \sum_{m=-\infty}^{\infty} \bar{g}_{k-2m} v_{j,m} + \sum_{m=-\infty}^{\infty} \bar{h}_{k-2m} w_{j,m}, \tag{483a}$$

其中在得出最后等式时我们用到了练习[464]和[479]的结果．注意在最后一行第一部分求和时不能表示为$\{\bar{g}_m\}$滤波$\{v_{j,m}\}$，因为当 k 为偶数时，求和只用到了$\{\bar{g}_m\}$的偶数角标部分的值，而当 k 为奇数时，求和只用到了$\{\bar{g}_m\}$的奇数角标部分的值(对包括$\{\bar{h}_m\}$求和类似的结论也成立)．

11.9 消失矩

令$\mathcal{N}_m$表示小波函数 $\psi(\cdot)$的 m 阶矩：

$$\mathcal{N}_m \equiv \int_{-\infty}^{\infty} t^m \psi(t)\,\mathrm{d}t. \tag{483b}$$

本节，我们探讨 r 阶消失矩对 $\psi(\cdot)$的影响，r 阶消失矩的意思是$\mathcal{N}_m=0$，对 $m=0$，1，…，$r-1$ 成立，其中 r 为一正整数．消失矩的意义源于如下事实(在评论与扩展第[1]条中简单讨论)，如果 $\psi(\cdot)$和它的 $r-1$ 次微分处处连续且满足某个有界性条件，则 $\psi(\cdot)$有 r 阶消失矩．$\psi(\cdot)$的连续性和足够阶可微性帮助消除了信号分析中由于小波函数本身的人为误差——通俗地说，$\psi(\cdot)$的消失矩的阶越高，从一个逼近空间到另一个逼近空间的转换越光滑．现在，只是如果$\psi(\cdot)$有 r 阶消失矩，且 $\psi(\cdot)$和它的 $r-1$ 阶导数都连续．这就是 Daubechies(1993)将消失矩条件合并到他的滤波器设计中的动机，像在本节讨论的．值得注意的是，具有 r 阶消失矩的小波 $\psi(\cdot)$不保证有 $r-1$ 阶连续导数，即"当且仅当"不成立，仅仅是"仅当"(我们转到评论与扩展第[2]条)．

加在 $\psi(\cdot)$的消失矩条件相当于加在它的傅里叶变换 $\Psi(\cdot)$的条件．如果令$\Psi^{(m)}(\cdot)$表示 $\Psi(\cdot)$的第 m 阶导数，那么我们有

$$\Psi^{(m)}(f) = (-\mathrm{i}2\pi)^m \int_{-\infty}^{\infty} t^m \psi(t) \mathrm{e}^{-\mathrm{i}2\pi ft}\,\mathrm{d}t,$$

由此可以看出$\mathcal{N}_m=0$ 意味着 $\Psi^{(m)}(0)=0$ ；即加在 $\psi(\cdot)$消失矩条件意味着 $\psi(\cdot)$的傅里叶变换及其一定阶的导数在频率为 0 的时候等于 0．

消失矩条件同时也隐含 $\bar{H}(\cdot)$上的某个条件，这是在式(472)中用来定义 $\psi(\cdot)$的序列$\{\bar{h}_l\}$的傅里叶变换．为此，对式(476c)(式(472)中频域)进行微分得

$$\Psi^{(1)}(f) = \frac{\Phi^{(1)}\left(\frac{f}{2}\right)}{2^{3/2}}\bar{H}\left(\frac{f}{2}\right) + \frac{\Phi\left(\frac{f}{2}\right)}{2^{3/2}}\bar{H}^{(1)}\left(\frac{f}{2}\right).$$

因为 $\bar{H}(0)=0$ (参见式(476b))和因为 $\Phi(0)=1$ (参见式(468b))，我们可以看到 $\Psi^{(1)}(0)=0$ 意味着 $\bar{H}^{(1)}(0)=0$．

练习[484] 通过证明，对任意的 $m=0, 1, \cdots, r-1$，$\Psi^{(m)}(0)=0$ 意味着 $\bar{H}^{(m)}(0)=0$，来推广上述结论. ◁

因此，如果 $\psi(\cdot)$ 有 r 阶消失矩，则对任意的 $m=0, 1, \cdots, r-1$，$\bar{H}^{(m)}(0)=0$. 由于

$$\bar{H}^{(m)}(0)=(-\mathrm{i}2\pi)^m\sum_{l=-\infty}^{\infty} l^m \bar{h}_l,$$

所以 $\psi(\cdot)$ 有 r 消失矩意味着序列 $\{\bar{h}_l\}$ 也有 r 阶离散的消失矩.

现在 $\{\bar{h}_l\}$ 是相应于 $\{\bar{g}_l\}$ 的正交镜像滤波器，所以通过式(476a)，加在 $\bar{H}(\cdot)$ 的条件变为加在 $\bar{G}(\cdot)$ 的条件. 首先注意

$$\bar{H}\left(\frac{1}{2}\right)=-\mathrm{e}^{\mathrm{i}\pi(L-1)}\bar{G}(0)=\bar{G}(0)=\sqrt{2},\quad \bar{H}(0)=-\bar{G}\left(\frac{1}{2}\right)=0$$

(使用式(466b)和式(476b)，并且回忆起 L 为偶数). 对式(476a)的左右两端进行微分，并且令 $f=0$，我们得到

$$\bar{H}^{(1)}(0)=-\mathrm{i}2\pi(L-1)\bar{G}\left(\frac{1}{2}\right)+\bar{G}^{(1)}\left(\frac{1}{2}\right)=\bar{G}^{(1)}\left(\frac{1}{2}\right),$$

所以条件 $\bar{H}^{(1)}(0)=0$ 意味着 $\bar{G}^{(1)}\left(\frac{1}{2}\right)=0$. 进而微分显示，对于 $m=0, 1, \cdots, r-1$，$\bar{H}^{(m)}(0)=0$ 意味着对于 $m=0, 1, \cdots, r-1$，$\bar{G}^{(m)}\left(\frac{1}{2}\right)=0$. 因此，$\psi(\cdot)$ 有 r 阶消失矩也意味着对任意的 $m=0, 1, \cdots, r-1$，$\bar{G}^{(m)}\left(\frac{1}{2}\right)=0$.

对于复值变量 z，让我们定义

$$\bar{\bar{G}}(z)\equiv\sum_{l=-\infty}^{\infty}\bar{g}_l z^{-l},$$

这称为序列 $\{\bar{g}_l\}$ 的 z 变换(Oppenheim and Schafer, 1989, p. 149). 如果令 $z=\mathrm{e}^{\mathrm{i}2\pi f}$，我们得到

$$\bar{\bar{G}}(\mathrm{e}^{\mathrm{i}2\pi f})=\sum_{l=-\infty}^{\infty}\bar{g}_l\mathrm{e}^{-\mathrm{i}2\pi fl}\equiv\bar{G}(f),$$

即 $\{\bar{g}_l\}$ 的傅里叶变换(实值变量 f 的函数). 回顾一下对 $\phi(\cdot)$ 和 $\psi(\cdot)$ 在区间 $(-(L-1), 0]$ 的有限支撑相当于 $\{\bar{h}_l\}$ 和 $\{\bar{g}_l\}$ 在 $-(L-1)\leqslant l\leqslant 0$ 之外为零. 所以，$\bar{\bar{G}}(\cdot)$ 是 z 非负指数的 $L-1$ 次多项式，我们可以写成

$$\bar{\bar{G}}(z)=\sum_{l=-(L-1)}^{0}\bar{g}_l z^{-l}=\bar{g}_{-(L-1)}\prod_{l=1}^{L-1}(z-z_l),\tag{484}$$

其中 z_l 是多项式 $\bar{\bar{G}}(\cdot)$ 除以 $\bar{g}_{-(L-1)}$ 后的第 l 个根(假定我们能够取 $\bar{g}_{-(L-1)}\neq 0$). 因为

$$\bar{G}(f)=\bar{\bar{G}}(\mathrm{e}^{\mathrm{i}2\pi f})=\bar{g}_{-(L-1)}\prod_{l=1}^{L-1}(\mathrm{e}^{\mathrm{i}2\pi f}-z_l),$$

由 $e^{i2\pi f}$，我们可以将 $\bar{G}(\cdot)$视为 $L-1$ 次三角多项式.

假设 $\psi(\cdot)$有 r 阶消失矩，从而当 $m=0, 1, \cdots, r-1$ 时有 $\bar{G}^{(m)}\left(\frac{1}{2}\right)=0$. 由于 $\bar{G}^{(m)}\left(\frac{1}{2}\right)=\bar{\bar{G}}^{(m)}(e^{i\pi})=\bar{\bar{G}}^{(m)}(-1)$，由此得到 $\bar{\bar{G}}^{(m)}(-1)=0$. 我们认为加在 $\bar{\bar{G}}(\cdot)$的导数上的这个条件意味着 $\bar{\bar{G}}(\cdot)$的 r 个根，计为 $z_1, \cdots, z_r$，都等于-1. 为此，我们在 $z=-1$ 处对 $\bar{\bar{G}}(\cdot)$进行泰勒(Taylor)展开得(利用对所有的 $l\geq L$，$\bar{\bar{G}}^{(l)}(z)=0$，因为 $\bar{\bar{G}}(\cdot)$是 $L-1$ 次多项式这一事实)

$$\bar{\bar{G}}(z)=\sum_{l=0}^{L-1}\frac{(z+1)^l}{l!}\bar{\bar{G}}^{(l)}(-1)=\sum_{l=r}^{L-1}\frac{(z+1)^l}{l!}\bar{\bar{G}}^{(l)}(-1)$$

$$=(z+1)^r\sum_{l=r}^{L-1}\frac{(z+1)^{l-r}}{l!}\bar{\bar{G}}^{(l)}(-1).$$

与式(484)联合产生

$$\bar{\bar{G}}(z)=\bar{g}_{-(L-1)}(z+1)^r\prod_{l=r+1}^{L-1}(z-z_l). \tag{485a}$$

令 $z=e^{i2\pi f}$，给出上式在频域中等价于

$$\bar{G}(f)=\bar{\bar{G}}(e^{i2\pi f})=\bar{g}_{-(L-1)}(e^{i2\pi f}+1)^r\prod_{l=r+1}^{L-1}(e^{i2\pi f}-z_l)=\left(\frac{e^{i2\pi f}+1}{2}\right)^r Q(f), \tag{485b}$$

其中

$$Q(f)\equiv 2^r\,\bar{g}_{-(L-1)}\prod_{l=r+1}^{L-1}(e^{i2\pi f}-z_l).$$

注意到因为 $\bar{G}(0)=\sqrt{2}$，则还必有 $Q(0)=\sqrt{2}$. 另外，因为 $z_l\neq -1$，对于 $l=r+1, \cdots, L-1$ 成立，我们知道 $Q\left(\frac{1}{2}\right)\neq 0$. 很容易看到 $Q(\cdot)$是有单位周期的 $L-r-1$ 次的三角多项式.

构造有有限支撑的小波的关键是式(485b). 定义$\bar{\mathcal{G}}(f)\equiv|\bar{G}(f)|^2$是平方增益函数，使

$$\bar{\mathcal{G}}(f)=\left|\frac{e^{i2\pi f}+1}{2}\right|^{2r}|Q(f)|^2=\cos^{2r}(\pi f)\,\mathcal{Q}(f),$$

其中$\mathcal{Q}(f)\equiv|Q(f)|^2$.

练习[485] 证明，因为序列$\{\bar{g}_l\}$是实值的，所以$\bar{\mathcal{G}}(f)$和$\mathcal{Q}(f)$都是 $\cos(2\pi f)$的多项式. ◁

因为 $\cos(2\pi f)=1-2\sin^2(\pi f)$，我们可以将$\mathcal{Q}(f)$写成 $\sin^2(\pi f)$的多项式，因此

$$\bar{\mathcal{G}}(f)=\cos^{2r}(\pi f)\,\mathcal{P}(\sin^2(\pi f)),$$

其中$\mathcal{P}(\cdot)$是待定多项式. 借助于$\bar{\mathcal{G}}(\cdot)$，式(467b)的正交性条件可以写成

$$\bar{\mathcal{G}}(f)+\bar{\mathcal{G}}\left(f+\frac{1}{2}\right)=2,$$

即

$$\cos^{2r}(\pi f)\ \mathcal{P}(\sin^2(\pi f))+\sin^{2r}(\pi f)\ \mathcal{P}(\cos^2(\pi f))=2$$

或者

$$(1-y)^r\ \mathcal{P}(y)+y^r\ \mathcal{P}(1-y)=2,$$

对任意的 $y\in[0,1]$ 成立. Daubechies(1992，命题 6.12，p. 171)证明了$\mathcal{P}(\cdot)$必须有如下形式：

$$\mathcal{P}(y)=2\left[\mathcal{P}_r(y)+y^r\ \mathcal{R}\left(\frac{1}{2}-y\right)\right],$$

其中

$$\mathcal{P}_r(y)\equiv\sum_{k=0}^{r-1}\binom{r-1+k}{k}y^k,$$

$\mathcal{R}(\cdot)$是一奇次多项式且选得使 $y\in[0,1]$ 时$\mathcal{P}(y)\geqslant 0$. 由此，$\mathcal{Q}(\cdot)$的形式完全知道了，但是，为了用式(485b)来得到滤波器$\{\bar{g}_l\}$，我们实际上需要知道$\mathcal{Q}(f)$的“平方根”$Q(f)$，求$\mathcal{Q}(f)=\mathcal{P}(\sin^2(\pi f))$的“平方根”在数学中叫谱分解.

Daubechies(1988)选择$\mathcal{R}(y)\equiv 0$ 构造了一族尺度滤波器，$\mathcal{P}(\cdot)\equiv 2\mathcal{P}_r(\cdot)$且平方增益函数为

$$\bar{\mathcal{G}}(f)=2\cos^{2r}(\pi f)\sum_{k=0}^{r-1}\binom{r-1+k}{k}\sin^{2k}(\pi f),\tag{486}$$

这与式(105a)取 $r=\frac{L}{2}$时是恒等的. 在下节，我们取 $L=2$，4，6 和 8 来对上面的谱分解进行讨论.

11.9 节的评论与扩展

[1]令 C^n表示使$\gamma^{(0)}(\cdot)$，$\gamma^{(1)}(\cdot)$，…，$\gamma^{(n)}(\cdot)$处处连续的$\gamma(\cdot)$函数集，其中$\gamma^{(k)}(\cdot)$表示 $\gamma(\cdot)$的第 k 阶导数(定义 $\gamma^{(0)}(\cdot)=\gamma(\cdot)$). 令 $\psi(\cdot)$是一个函数，使得$\{\psi_{j,k}(\cdot):k\in\mathbf{Z}\}$构成 $L^2(\mathbf{R})$中函数的一个规范正交基(如果 $\psi(\cdot)$为一小波函数，这能满足). Daubechies(1992，推论 5.5.2，p. 154)证明了：如果 $\psi(\cdot)\in C^{r-1}$，且如果 $\psi^{(k)}(\cdot)$对于 $k\leqslant r-1$ 是有界的，并且如果

$$|\psi(t)|\leqslant\frac{c}{(1+|t|)^{r+\epsilon}},$$

其中 c 是一常数，且 $\epsilon>0$，则 $\psi(\cdot)$具有 r 阶消失矩(即在式(483b)中，$m=0$，1，…，$r-1$ 时都有$\mathcal{N}_m=0$).

[2]尺度和小波函数的光滑性的一种度量通过它们所谓的正则性给出，我们可以通过利普希茨(Lipschitz)(或者赫尔德(Hölder)条件)定义. 如果对任意的 t，$t'\in\mathbf{R}$ 有

$$|\gamma(t)-\gamma(t+t')|<c|t'|^{\beta},$$

则称函数 $\gamma(\cdot)$ 有 β 阶 Lipschitz 正则阶，其中 $0<\beta\leqslant 1$，c 是一常数(实际中，认为 t' 很小)．由 $\gamma(\cdot)$ 的 n 阶导数代替 $\gamma(\cdot)$ 可以得到高阶的正则性 $\alpha=n+\beta$，这样定义 α 阶 Hölder 空间介于 C^n 和 C^{n+1} 之间(回顾 C^n 表示 $\gamma(\cdot)$ 函数集使得 $\gamma(\cdot)$ 和它的 1 到 n 阶导数都处处连续)．在 α 阶 Hölder 空间中的函数称为是 α 阶(赫尔德)正则．

Rioul(1992)估计了对 $L=4$，6 和 8 时 D(L)，$\phi(\cdot)$ 和 $\psi(\cdot)$ 函数的赫尔德正则性，分别如下：

$$\phi(\cdot),\psi(\cdot)\in C^{0.5500},\quad 即\ \phi(\cdot),\psi(\cdot)\in C^{0},\notin C^{1};$$
$$\phi(\cdot),\psi(\cdot)\in C^{1.0878},\quad 即\ \phi(\cdot),\psi(\cdot)\in C^{1},\notin C^{2};$$
$$\phi(\cdot),\psi(\cdot)\in C^{1.6179},\quad 即\ \phi(\cdot),\psi(\cdot)\in C^{1},\notin C^{2},$$

其中，例如 D(4)，$\phi(\cdot)$ 和 $\psi(\cdot)$ 函数连续，因此属于 C^0，但是它们不属于 C^1(即它们没有连续导数)．注意既然 $\psi(\cdot)$ 是 $\phi(2t)$ 的平移的有限线性组合，所以认为 $\phi(\cdot)$ 和 $\psi(\cdot)$ 有相同的正则阶是合理的．

对于比较大的 L 的 D(L)族，Daubechies(1992)发现近似地有

$$\phi(\cdot),\psi(\cdot)\in C^{L/10},\quad 即\ \phi(\cdot),\psi(\cdot)\in C^{r/5},$$

因此大约 80%的零矩量是“没用”的，即同样的正则性理论上只有 $L/10$ 消失矩可以得到．这意味着 $\psi(\cdot)$ 有 r 阶消失矩不能保证有 $\psi(\cdot)\in C^{r-1}$．

[3]回想 $\psi(\cdot)$ 有 r 阶消失矩意味着

$$\tilde{G}^{(m)}\left(\frac{1}{2}\right)=\tilde{G}^{(m)}\left(-\frac{1}{2}\right)=0,m=0,\cdots,r-1;$$

由于 $\tilde{G}(f)=G(-f)$，所以有 $G^{(m)}\left(\frac{1}{2}\right)=0$．对于 $\{g_l\}$ 的增益函数 $|G(\cdot)|$ 在 $f=\frac{1}{2}$ 处随着 r 的增大而变得更扁．Daubechies(1992，p. 245－7)指出：对包括 $G(\cdot)$ 的多重应用的级联的滤波器进行隔行下采样(应用于离散小波变换)，对于等价滤波器的增益函数在高频处将出现波纹，除非 $|G(\cdot)|$ 在 $f=\frac{1}{2}$ 处非常扁平．

11.10 谱分解和滤波器系数

对于式(486)中的 $\bar{\mathcal{G}}(\cdot)$ 的分解，我们先回顾一下实系数滤波器(即实值脉冲响应序列)的延时、相位以及属性的一些结果．我们的出发点是式(485a)取 $r=L/2$．令

$$\tilde{G}(z)=\tilde{g}_{-(L-1)}(z+1)^{L/2}\bar{\tilde{Q}}_0(z),\tag{487}$$

对此我们定义

$$\bar{\tilde{Q}}_0(z)\equiv\prod_{l'=\frac{L}{2}+1}^{L-1}(z-z_{l'})=\prod_{l=1}^{\frac{L}{2}-1}(z-z_l);$$

右边等式仅仅是为了脚标的方便．注意到$\bar{\bar{Q}}_0(z)$是z的$\frac{L}{2}-1$次非负指数多项式；$z^{\frac{L}{2}-1}$的系数为1；且它所有的$\frac{L}{2}$个系数都一定是实数(因为$\bar{G}(z)$是实值滤波器的z变换)．写$z=e^{i2\pi f}$，上式变为一个传递函数，因此我们可以将上式看为一个长度为$\frac{L}{2}$的滤波器的z变换．因为两滤波器的卷积等价于它们的z变换的乘积，滤波器的上述z变换可解释为两个$\frac{L}{2}-1$滤波器的卷积．长度为2的滤波器也常称为“偶极子”，以后我们采用这个术语．第l个偶极子具有z变换$z-z_l \equiv z+b_l$，$e^{i2\pi f}+b_l$的传递函数和在$t=-1$的等于1的以及$t=0$时$b_l=-z_l$其他为零的脉冲响应序列．让我们写这个偶极子的脉冲响应序列为$\{1,\ \boldsymbol{b}_l\}$，其中将指标$t=0$的系数用粗体表示．注意为了得到长度为$\frac{L}{2}$的实值系数的滤波器，我们可以用偶极子$\{1,\ \boldsymbol{b}_l\}$，而b_l为实数或者具有系数为共轭复数的数对，我们说$\{1,\ \boldsymbol{b}_m\}$和$\{1,\ \boldsymbol{b}_n\}$是复共轭对，如果$b_n=b_m^*$．

现在，让我们比较第l个偶极子的两个非零系数的大小，即我们想比较$|b_l|=|z_l|$与1．假如$|b_l|>1$，在此情形下，根z_l位于复平面单位圆外面．因为两个非负系数的偶极子中有一个在很长时间以后才出现，这样的偶极子称为**最大延时或者最大相位偶极子**．如果全部$\frac{L}{2}-1$个根z_l都在单位圆外，则宽度为$\frac{L}{2}$的实系数滤波器是$\frac{L}{2}-1$最大延时偶极子的卷积，且得到的滤波器被描述为最大延时或者最大相位滤波器．(参见例如，Robinson，1980；Robinson and Treitel，1980)．它的(在4.8节讨论的)部分能量序列的建立是任一具有同样平方增益函数同宽度滤波器中最慢的．

下面让我们考虑偶极子$\{\boldsymbol{b}_l,\ 1\}$，即它的脉冲响应序列当$t=0$时等于b_l，$t=1$时等于1，其他为0(这相当于时间逆转$\{1,\ \boldsymbol{b}_l\}$)．如果仍然有$|b_l|>1$，则这两个非零系数中有一个在早些时候就变大．$\{\boldsymbol{b}_l,\ 1\}$的z变换为b_l+z^{-1}，它有一个根$-1/b_l=1/z_l$位于单位圆内．这样的偶极子被称为**最小延时或者最小相位偶极子**．如果$\{\boldsymbol{b}_l,\ 1\}$，$l=1,\ \cdots,\ \frac{L}{2}-1$，都是最小延时偶极子，则通过它们的卷积得到的宽为$\frac{L}{2}$的实值滤波器可以描述成最小延时或者最小相位滤波器．它的部分能量序列的建立是任何同宽度滤波器的平方增益函数中最快的．

现在，对z变换是$\bar{\bar{Q}}_0(z)$的滤波器的平方增益函数是通过在$z=e^{i2\pi f}$处估计$|\bar{\bar{Q}}_0(z)|^2$获得的．

练习[488]　证明当$z=e^{i2\pi f}$时，我们有

$$[\bar{\bar{Q}}_0(z)]^*=\bar{\bar{Q}}_0(z^{-1}),\text{从而有}|\bar{\bar{Q}}_0(z)|^2=\bar{\bar{Q}}_0(z)\bar{\bar{Q}}_0(z^{-1}).$$ ◁

现在$\bar{\bar{Q}}_0(z)$是偶极子的子变换项的乘积．具有实值系数的偶极子$\{1,\ \boldsymbol{b}_l\}$贡献$(z+b_l)(z^{-1}+b_l)$到$|\bar{\bar{Q}}_0(z)|^2$．注意$|b_l|>1$相应于一个最大延时偶极子$\{1,\ \boldsymbol{b}_l\}$和一个最小延时偶极子$\{\boldsymbol{b}_l,\ 1\}$的卷积：由于它是滤波器与它时间逆的卷

积，所以这个结果仅仅是最大延时偶极子的自相关. 一对系数互为复共轭的偶极子$\{1, b_m\}$和$\{1, b_m^*\}$将贡献$(z+b_m)(z+b_m^*)(z^{-1}+b_m)(z^{-1}+b_m^*)$到$|\bar{Q}_0(z)|^2$. 因为自相关是$\{1, b_m\}$的复共轭与它的时间逆的卷积，所以$(z+b_m^*)(z^{-1}+b_m)$是$\{1, b_m\}$自相关偶极子的 z 变换. $|b_m|>1$ 相应于一个最大延时偶极子$\{1, b_m\}$和一个最小延时偶极子$\{b_m, 1\}$的卷积，还有$(z+b_m)(z^{-1}+b_m^*)$是$\{1, b_m^*\}$的自相关的 z 变换，也相应于一个最大延时偶极子和一个最小延时偶极子的卷积.

下面我们从$|\bar{Q}_0(z)|^2$着手来举一简单的例子. 有很多方法可以得到$\bar{\bar{Q}}_0(z)$：我们可以取所有的最大延时偶极子，或所有的最小延时偶极子，或具有实值系数的最大延时偶极子和最小延时偶极子的混合，以及具有复值系数的同为最大延时或最小延时的偶极子对. 从具有$\frac{L}{2}-1$ 个实值系数的偶极子的集合开始，有 $2^{\frac{L}{2}-1}$ 个长度为$\frac{L}{2}$的实值滤波器且具有同样数量级的 z 变换的平方(或平方增益函数)，但相位函数不同. 如果具有实值系数的 K_0 个偶极子和复值共轭系数的 K_1 个偶极子对做卷积，有 $K_0+2K_1=\frac{L}{2}-1$，则有 $2^{K_0+K_1}$ 个长度为$\frac{L}{2}$的实值滤波器，具有同样数量的 z 变换的平方(或平方增益函数)，但相位函数不同.

- **一阶消失矩与哈尔尺度滤波器**

当 $L=2$ 时有 $r=L/2=1$ 阶消失矩，我们有$\mathcal{P}(\sin^2(\pi f))\equiv 2\ \mathcal{P}_1(\sin^2(\pi f))=2$. 因此

$$\bar{\mathcal{G}}(f)=|\bar{G}(f)|^2=2\cos^2(\pi f)=2\left[\frac{1}{2}+\frac{1}{2}\cos(2\pi f)\right]=2\left[\frac{1}{2}+\frac{1}{4}(e^{-i2\pi f}+e^{i2\pi f})\right].$$

从而

$$|\bar{G}(z)|^2=2\left[\frac{1}{4}(z^{-1}+2+z)\right]=\frac{1}{2}(z+1)(z^{-1}+1).$$

取 $\bar{G}(z)=\sum_{l=-1}^{0}\bar{g}_l z^{-l}=\bar{g}_{-1}z+\bar{g}_0=\pm(z+1)/\sqrt{2}$，注意我们期望在 $z=-1$ 处有一个单根$(r=1)$且$\bar{G}(z)\bar{G}(z^{-1})=|\bar{G}(z)|^2$. 为了$\bar{G}(1)=\sqrt{2}$(即$G(0)=\sqrt{2}$)，我们取正号. 相应的滤波器的系数是$\bar{g}_{-1}=\bar{g}_0=1/\sqrt{2}$. 哈尔尺度滤波器是相应于时间逆的滤波器，即$\{g_l=\bar{g}_{-l}\}$，因此 $g_0=g_1=1/\sqrt{2}$，如式(463d).

- **二阶消失矩与 $D(4)$ 尺度滤波器**

当 $L=4$(即 $r=2$ 阶消失矩)时，我们有$\mathcal{P}(\sin^2(\pi f))\equiv 2\ \mathcal{P}_2(\sin^2(\pi f))=2+4\sin^2(\pi f)$. 因此

$$\mathcal{G}(f)=\cos^4(\pi f)[2+4\sin^2(\pi f)]=\left[\frac{1}{2}+\frac{1}{2}\cos(2\pi f)\right]^2[4-2\cos(2\pi f)].$$

从而

$$|\bar{G}(z)|^2=\frac{1}{16}(z^{-1}+2+z)^2(-z^{-1}+4-z).$$

我们从一阶消失矩的情况知道，$(z^{-1}+2+z)^2$可以分解为$(z+1)^2(z^{-1}+1)^2$. 我们只需要另外分解$(-z^{-1}+4-z)$. 如果我们乘以$-z$，就得到 z^2-4z+1，它有互为倒数的两个实根，$2+\sqrt{3}\equiv z_1$和 $2-\sqrt{3}\equiv z_1^{-1}$；即

$$-z^{-1}+4-z=\frac{z^2-4z+1}{-z}=\frac{(z-z_1)(z-z_1^{-1})}{-z}=z_1^{-1}(z-z_1)(z^{-1}-z_1).$$

令 $b_1=-z_1$，注意到$|b_1|>1$，因此我们有

$$|\overline{G}(z)|^2=C^2(z+1)^2(z^{-1}+1)^2(z+b_1)(z^{-1}+b_1),$$

其中 $C^2\equiv z_1^{-1}/16$ 从而 $C=\pm\frac{1-\sqrt{3}}{4\sqrt{2}}$. 如果我们继续像在 $r=1$ 的情形，用包含因子 $z+1$ 的项来构造 $\overline{G}(z)$且取 $C<0$，我们有两种可能的$\overline{G}(z)$的定义，即

$$\overline{G}(z)\equiv C(z+1)^2(z+b_1),$$

因而

$$\overline{G}(z^{-1})=C(z^{-1}+1)^2(z^{-1}+b_1)$$

或者

$$\overline{G}(z)\equiv Cz(z+1)^2(z^{-1}+b_1)=b_1C(z+1)^2(z+b_1^{-1}), \tag{490a}$$

因而

$$\overline{G}(z^{-1})=Cz^{-1}(z^{-1}+1)^2(z+b_1),$$

这两种定义都导出了需要的性质，$\overline{G}(z)$是 z 的非负指数次幂的多项式，且 $\overline{G}(z)\overline{G}(z^{-1})=|\overline{G}(z)|^2$和 $\overline{G}(1)=\sqrt{2}$(注意也可以是 $\overline{G}(z)$的另外一种定义，期望它在 $z=-1$ 处有一重根). 第一种定义利用最大延时偶极子$\{1, b_1\}$，第二种利用最小延时偶极子$\{b_1, 1\}$. 比较第一种定义与式(487)，我们注意到$\overline{Q}_0(z)=z+b_1=z-2-\sqrt{3}$，展开第一种定义，得到

$$\begin{aligned}\overline{G}(z)&=\frac{1-\sqrt{3}}{4\sqrt{2}}(z^3+[2+b_1]z^2+(1+2b_1)z+b_1) \qquad\qquad (490\text{b})\\&=\frac{1-\sqrt{3}}{4\sqrt{2}}z^3+\frac{3-\sqrt{3}}{4\sqrt{2}}z^2+\frac{3+\sqrt{3}}{4\sqrt{2}}z+\frac{1+\sqrt{3}}{4\sqrt{2}},\end{aligned}$$

从而系数为

$$\{\overline{g}_{-3},\overline{g}_{-2},\overline{g}_{-1},\overline{g}_0\}=\left\{\frac{1-\sqrt{3}}{4\sqrt{2}},\frac{3-\sqrt{3}}{4\sqrt{2}},\frac{3+\sqrt{3}}{4\sqrt{2}},\frac{1+\sqrt{3}}{4\sqrt{2}}\right\}.$$

滤波器$\{g_l\}$的系数用时间逆得出，即

$$\{g_0,g_1,g_2,g_3\}=\left\{\frac{1+\sqrt{3}}{4\sqrt{2}},\frac{3+\sqrt{3}}{4\sqrt{2}},\frac{3-\sqrt{3}}{4\sqrt{2}},\frac{1-\sqrt{3}}{4\sqrt{2}}\right\},$$

这就是在式(75d)中所述的 D(4)尺度滤波器.

练习[490] 证明 $\overline{G}(z)$的第二种定义(即在式(490a)中的一个包含一个最小延时偶极子)导出

$$\{g_0,g_1,g_2,g_3\}=\left\{\frac{1-\sqrt{3}}{4\sqrt{2}},\frac{3-\sqrt{3}}{4\sqrt{2}},\frac{3+\sqrt{3}}{4\sqrt{2}},\frac{1+\sqrt{3}}{4\sqrt{2}}\right\},$$

与 D(4)尺度滤波器的系数是相反顺序. ◁

• 三阶消失矩与 D(6) 尺度滤波器

当 $L=6$ 从而 $r=3$ 时，我们有

$$\mathcal{P}(\sin^2(\pi f)) \equiv 2\,\mathcal{P}_3(\sin^2(\pi f)) = 2 + 6\sin^2(\pi f) + 12\sin^4(\pi f).$$

因此

$$\begin{aligned}\bar{\mathcal{G}}(f) &= \cos^6(\pi f)[2 + 6\sin^2(\pi f) + 12\sin^4(\pi f)] \\ &= \left[\frac{1}{2} + \frac{1}{2}\cos(2\pi f)\right]^3 [8 - 9\cos(2\pi f) + 3\cos^2(2\pi f)].\end{aligned}$$

用 $(z+z^{-1})/2$ 取代上式的 $\cos(2\pi f)$ 得到

$$|\bar{G}(z)|^2 = \frac{1}{64}(z+1)^3(z^{-1}+1)^3\,\frac{1}{4}(3z^{-2} - 18z^{-1} + 38 - 18z + 3z^2).$$

我们可以利用 S-PLUS 函数多项式求根来求出

$$1 - 6z + \frac{38}{3}z^2 - 6z^3 + z^4 = (z - z_1)(z - z_1^*)(z - z_2)(z - z_2^*),$$

其中

$$z_1 \doteq 2.712\,748\,621\,955\,979\,5 + 1.443\,886\,782\,618\,004\,0\mathrm{i} \qquad (491)$$

和

$$z_2 \doteq 0.287\,251\,378\,044\,020\,9 + 0.152\,892\,333\,882\,199\,2\mathrm{i} \doteq \frac{1}{z_1^*} = \frac{z_1}{|z_1|^2}.$$

所以我们可以写成

$$|\bar{G}(z)|^2 = \frac{3}{256|z_1|^2}(z+1)^3(z^{-1}+1)^3(z-z_1)(z-z_1^*)(z^{-1}-z_1)(z^{-1}-z_1^*),$$

其中，注意因为 $|z_1|>1$，所以项 $z-z_1$ 和 $z-z_1^*$ 相应于最大延时偶极子 $\{1, -z_1\}$ 和 $\{1, -z_1^*\}$. 如果再用含有因子 $z+1$ 的项和最大延时偶极子来构造 $\bar{G}(z)$，则又可以得到

$$\bar{G}(z) = \frac{\sqrt{3}}{16|z_1|}(z+1)^3(z-z_1)(z-z_1^*),$$

这是 z 的非负指数的多项式，且满足要求的条件 $\bar{G}(z)\bar{G}(z^{-1}) = |\bar{G}(z)|^2$ 和 $\bar{G}(1)=\sqrt{2}$. 我们比较上式与式(487)，注意 $\bar{Q}_0(z)=(z-z_1)(z-z_1^*)$. 展开上式并与 $\bar{g}_{-5}z^5 + \bar{g}_{-4}z^4 + \cdots + \bar{g}_0$ 相等得到

$$\begin{bmatrix}\bar{g}_{-5}\\ \bar{g}_{-4}\\ \bar{g}_{-3}\\ \bar{g}_{-2}\\ \bar{g}_{-1}\\ \bar{g}_0\end{bmatrix} = \frac{\sqrt{3}}{16|z_1|}\begin{bmatrix}1\\ 3-2r_1\\ 3-6r_1+|z_1|^2\\ 1-6r_1+3|z_1|^2\\ -2r_1+3|z_1|^2\\ |z_1|^2\end{bmatrix},$$

其中 r_1 为 z_1 的实部. 回想 $g_l = \bar{g}_{-l}$，计算给出如表 109 的对于 D(6) 尺度滤波器

的值.

• D(L)滤波器的一般形式

对于一般的 $L\geqslant 2$，我们可以使用式(487)和练习[488]给出(设 $z=\mathrm{e}^{\mathrm{i}2\pi f}$)

$$|\bar{G}(z)|^2=\bar{g}^2_{-(L-1)}|z+1|^L|\bar{Q}_0(z)|^2=\bar{g}^2_{-(L-1)}|z+1|^L\bar{Q}_0(z)\bar{Q}_0(z^{-1}),$$

其中 $\bar{Q}_0(z)$ 是 z 的非负指数次幂的实系数 $\frac{L}{2}-1$ 次多项式. 对 $L=4$ 的情形，我们有 $\bar{Q}_0(z)=z-2-\sqrt{3}\equiv z-r_1$，其中 r_1 为一实根；对 $L=6$ 的情形，我们有 $\bar{Q}_0(z)=(z-z_1)(z-z_1^*)$，其中 z_1 为式(491)的复根. 对于一般的 $L\geqslant 4$，我们总可以写成

$$|\bar{Q}_0(z)|^2=\prod_{l=1}^{K_0}(z-r_l)(z^{-1}-r_l)\prod_{l=1}^{K_1}(z-z_l)(z-z_l^*)(z^{-1}-z_l)(z^{-1}-z_l^*),$$

其中 $K_0+2K_1=\frac{L}{2}-1$；r_1，…，r_{K_0} 是实根(没有一个等于或者小于零)；而 z_1，…，z_{K_1} 为复根(Daubechies，1992，p. 173；Chui，1997，p. 81). 而对 $|\bar{G}(z)|^2$ 的表达式是唯一的. 因为 $|\bar{G}(\mathrm{e}^{\mathrm{i}2\pi f})|^2=|\bar{G}(f)|^2=\bar{\mathcal{G}}(f)$，其中 $\mathcal{G}(f)$ 由式(486)给出. 对 $|\bar{Q}_0(z)|^2$ 的表达式却不唯一. 实际上，如果我们转化任意一个实根 r_l 或者任意一对复根 z_l，z_l^*，则 $|\bar{Q}_0(z)|^2$ 也将改变. 例如，因为

$$(z-r_l)(z^{-1}-r_l)=r_l^2(z-r_l^{-1})(z^{-1}-r_l^{-1}),$$

我们可写成

$$|\bar{Q}_0(z)|^2=|\bar{Q}_1(z)|^2\Big(\prod_{l=1}^{K_0}r_l^2\Big),$$

其中

$$|\bar{Q}_1(z)|^2\equiv\prod_{l=1}^{K_0}(z-r_l^{-1})(z^{-1}-r_l^{-1})\prod_{l=1}^{K_1}(z-z_l)(z-z_l^*)(z^{-1}-z_l)(z^{-1}-z_l^*),$$

由此我们可得

$$|\bar{G}(z)|^2=\Big(\bar{g}^2_{-(L-1)}\prod_{l=1}^{K_0}r_l^2\Big)|z+1|^L|\bar{Q}_1(z)|^2.$$

现在的 $|\bar{Q}_1(z)|^2$ 相当于原来的 $|\bar{Q}_0(z)|^2$. 所以 $\bar{Q}_0(z)$ 有 $2^{K_0+K_1}$ 种可能的选择却得到同样的平方增益函数，相位函数不同.

现在已经很清楚，给出 Daubechies D(L)滤波器 $\{g_l\}$ 的 $\{\bar{g}_l\}$ 的选择是取

$$\bar{Q}_0(z)=\prod_{l=1}^{K_0}(z-r_l)\prod_{l=1}^{K_1}(z-z_l)(z-z_l^*),$$

其中$|r_1|$，…，$|r_{K_c}|$，$|z_1|$，…，$|z_{K_1}|$都大于1，换句话说，$\tilde{Q}_0(z)$是由最大延时偶极子构造的，上面 $L=4$ 和 $L=6$ 就是例子. 滤波器$\{\bar{g}_l, l=-(L-1), \cdots, 0\}$因此由$\bar{\bar{Q}}_0(z)$中的最大延时偶极子与形如$\{1, 1\}$的$L/2$个的偶极子的卷积形成. 因此宽度为$L$（与形如$\{1, 1\}$的$L/2$个偶极子）的滤波器和增益函数$|G(\cdot)|$，即$\{\bar{g}_l\}$，它的部分能量序列的建立是最慢的，因此$\{\bar{g}_l\}$是最大延时滤波器. 另一方面，宽度为$L$（同样与形如$\{1, 1\}$的$L/2$的偶极子）的滤波器是$\{\bar{g}_l\}$的时间逆和增益函数$|\bar{G}(\cdot)|$，即为一个 Daubechies $D(L)$滤波器$\{g_l, l=-(L-1), \cdots, 0\}$，它的部分能量序列的建立是最快的，因此$\{\bar{g}_l\}$是最小延时滤波器.

最小延时滤波器的相位绝大部分从零分开，因此最小延时或者最小相位滤波器的设计在小波文献中(如：Daubechies，1992，p. 255)有时候也称为极值相位. 对于这最小相位滤波器需要“前载”，从而尺度函数是非常不对称的(唯一的例外是 $L=2$（哈尔)时，因为尺度函数关于 $t=-\frac{1}{2}$对称).

• **四阶消失矩与 LA(8)尺度滤波器**

在 4.8 节中已经指出，Daubechies 小波族中最接近对称的尺度滤波器通过选择不同的谱分解来产生；即比 D(L)滤波器的选择更好的对$\bar{\bar{Q}}_0(z)$相应的最大延时偶极子的项的选择，使得式(112a)的表达式中根的选择尽可能地小. 对于$L=8$，

$$\mathcal{P}(\sin^2(\pi f)) = 2\,\mathcal{P}_4(\sin^2(\pi f)) = 2 + 8\sin^2(\pi f) + 20\sin^4(\pi f) + 40\sin^6(\pi f),$$

由此可得

$$|\bar{\bar{Q}}_0(z)|^2 \propto z^{-3} - 8z^{-2} + 26.2z^{-1} - 41.6 + 26.2z - 8z^2 + z^3 \qquad (493a)$$

(验证上式是练习[11.5]的前半部分). 利用 S－PLUS 多项式求根程序，可以写成

$$|\bar{\bar{Q}}_0(z)|^2 = (z-r_1)(z^{-1}-r_1)(z-z_1)(z-z_1^*)(z^{-1}-z_1)(z^{-1}-z_1^*),$$

其中

$$r_1 \doteq 3.040\,660\,461\,647\,448\,0$$

$$z_1 \doteq 2.031\,135\,512\,091\,439\,0 + 1.738\,950\,807\,644\,819\,\mathrm{i},$$

因为$|r_1|>1$和$|z_1|>1$，式(493b)中的项 $z-r_1$、$z-z_1$和 $z-z_1^*$ 相应于最大延时偶极子.

取

$$\bar{\bar{G}}(z) = -\left[\frac{5}{2^{11}} r_1^{-1} |z_1|^{-2}\right]^{1/2} (z+1)^4 (z-r_1)(z-z_1)(z-z_1^*) \qquad (493b)$$

得到一个滤波器$\{\bar{g}_l\}$，经过时间逆后，给出表 109 中的 Daubechies“最小延时”D(8)滤波器$\{g_l\}$(参见练习[11.5]的第二部分). 对于这种选择，

$$\bar{\bar{Q}}_0(z) = (z-r_1)(z-z_1)(z-z_1^*).$$

然而，有四种可能的选择来构造 $\bar{G}(z)$：相应于根 $\{r_1, z_1, z_1^*\}$ 如上，$\{1/r_1, z_1, z_1^*\}$，$\{r_1, 1/z_1, 1/z_1^*\}$ 或者 $\{1/r_1, 1/z_1, 1/z_1^*\}$. 第二种根的选择给出

$$\bar{G}(z) = \left[\frac{5}{2^{11}} r_1 |z_1|^{-2}\right]^{1/2} (z+1)^4 (z - r_1^{-1})(z - z_1)(z - z_1^*),$$

并得到滤波器 $\{\bar{g}_l\}$ 经时间逆后，给出得到如表 109 中的 Daubechies"混合延时" LA(8)滤波器 $\{g_l\}$. 在此情形下，

$$\bar{Q}_0(z) = (z - r_1^{-1})(z - z_1)(z - z_1^*).$$

(第三和第四种描述给出式(112a)更大值的表达式，所以不选择作为 LA(8)滤波器.)

11.10 节的评论与扩展

[1]这里我们所考虑的偶极子都有一个系数等于 1，因此我们可以表明最大/最小延时条件借助于项中其他系数与 1 进行比较. 偶极子最一般的形式为 $\{a_l, b_l\}$，其中 a_l 有先于 b_l 的发生顺序，a_l 和 b_l 均可为复值. 这样，如果 $|a_l| < |b_l|$ 则这个偶极子称为最大延时偶极子，如果 $|a_l| > |b_l|$ 则称之为最小延时偶极子.

[2]对于 Daubechies 最接近对称的(LA(L))尺度函数和小波函数的赫尔德正则性低于最小延时 D(L)尺度和小波函数，详见 Rioul(1992，表 1). 因此，令人惊讶地发现，给定同样的增益函数，增加对称性将降低正则性.

11.11 小结

本章我们研究了连续时间小波和离散时间小波分析之间的关系. 连续时间小波分析处理定义域是整个实轴 $\mathbf{R}$ 的函数 $\gamma(\cdot)$(或者称为信号). 为了简单起见，我们假设 $\gamma(\cdot)$ 是实值的且属于 $L^2(\mathbf{R})$，$L^2(\mathbf{R})$ 是实数域 $\mathbf{R}$ 内平方可积的函数空间. 我们通过下式定义对 $\gamma(\cdot)$ 进行第 j 阶伸缩和第 k 阶平移：

$$\gamma_{j,k}(t) \equiv \frac{\gamma\left(\frac{t}{2^j} - k\right)}{\sqrt{2^j}}, \quad -\infty < t < \infty,$$

其中 $j, k \in \mathbf{Z}$(整数集).

在信号的小波分析中的主要概念是多分辨分析(MRA)，定义它是满足以下性质的闭子空间的序列 $V_j \subset L^2(\mathbf{R})$，$j \in \mathbf{Z}$. 子空间必须是嵌套的

$$\cdots \subset V_3 \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset \cdots,$$

同时它们的并和交的闭包必须满足

$$\overline{\bigcup_{j\in\mathbf{Z}} V_j} = L^2(\mathbf{R}) \text{ 和 } \bigcap_{j\in\mathbf{Z}} V_j = \{0\},$$

其中$\{0\}$表示仅包含零函数(即对所有的 $t\in\mathbf{R}$，该函数的值均为 0)的集合. V_0空间是使 $\gamma(\cdot)\in V_0$ 当且仅当 $\gamma_{0,k}(\cdot)=\gamma(t-k)\in V_0$，$k\in\mathbf{Z}$；另外，$\gamma(\cdot)\in V_0$ 当且仅当 $\gamma_{j,0}(\cdot)\in V_j$，$j\in\mathbf{Z}$. 最后，必须存在一个函数 $\phi(\cdot)\in L^2(\mathbf{R})$，使得它的整数平移 $\{\phi_{0,k}(\cdot):k\in\mathbf{Z}\}$ 形成闭子空间 V_0 的一个规范正交基(这就推出 $\{\phi_{j,k}(\cdot):k\in\mathbf{Z}\}$形成 V_j的一个规范正交基). 函数 $\phi(\cdot)$称为尺度函数.

在多分辨分析中的子空间 V_j被称为是一个逼近空间，与它相应的尺度为 $\lambda_j=2^j$. 给定一个多分辨分析，信号 $x(\cdot)\in V_0$可以表示为

$$x(t)=\sum_{k=-\infty}^{\infty}\langle x(\cdot),\phi_{0,k}(\cdot)\rangle\phi_{0,k}(t)=\sum_{k=-\infty}^{\infty}v_{0,k}\phi_{0,k}(t),$$

其中

$$v_{0,k}\equiv\langle x(\cdot),\phi_{0,k}(\cdot)\rangle=\int_{-\infty}^{\infty}x(t)\phi_{0,k}(t)\,\mathrm{d}t.$$

我们通过向子空间 $V_1\subset V_0$上投影可以得到 $x(\cdot)$的一个比较粗糙尺度上的逼近

$$s_1(t)=\sum_{k=-\infty}^{\infty}v_{1,k}\phi_{1,k}(t),$$

其中 $v_{1,k}\equiv\langle x(\cdot),\ \phi_{1,k}(\cdot)\rangle$. 函数 $s_1(\cdot)\in V_1$是对单位尺度信号 $x(\cdot)$的二尺度的粗糙逼近. $v_{1,k}$是二尺度系数. $v_{0,k}$与 $v_{1,k}$之间的关系由下式给出：

$$v_{1,k}=\sum_{m=-\infty}^{\infty}\bar{g}_{m-2k}v_{0,m}=\sum_{l=-\infty}^{\infty}g_l v_{0,2k-l},$$

其中 $\bar{g}_l\equiv\langle\phi_{0,l}(\cdot),\ \phi_{1,0}(\cdot)\rangle$，$g_l\equiv\bar{g}_{-l}$(这里的$\{g_l\}$与 4.3 节中所介绍的尺度滤波器一样，因此$\{\bar{g}_l\}$仅仅是它的时间逆). 这个方案可以推广：函数 $s_{j-1}(\cdot)\in V_{j-1}$ 通过向子空间 $V_j\subset V_{j-1}$上投影可以得到

$$s_j(t)\equiv\sum_{k=-\infty}^{\infty}v_{j,k}\phi_{j,k}(t),$$

$v_{j,k}\equiv\langle s_{j-1}(\cdot),\ \phi_{j,k}(\cdot)\rangle=\langle x(\cdot),\ \phi_{j,k}(\cdot)\rangle$是相应于尺度 $\lambda_j=2^j$的尺度系数. 这里 $s_j(\cdot)$是对信号 $x(\cdot)$的尺度 λ_j的逼近(也可视为对 $s_{j-1}(\cdot)$的逼近). $v_{j,k}$可以通过下式递推计算：

$$v_{j,k}=\sum_{m=-\infty}^{\infty}\bar{g}_{m-2k}v_{j-1,m}=\sum_{l=-\infty}^{\infty}g_l v_{j-1,2k-l}.$$

多分辨子空间的嵌套提升两尺度差分方程(465a)，即

$$\phi(t)=\sqrt{2}\sum_{l=-\infty}^{\infty}\bar{g}_l\phi(2t-l).$$

如果我们假设对于 l 在范围 $-(L-1), \cdots, 0$ 之外时 $\{\bar{g}_l\} \equiv \{g_{-l}\}$ 是零，则在区间 $(-(L-1), 0]$ 外 $\phi(\cdot)$ 是零，这时称 $\phi(\cdot)$ 有有限支撑．我们记这个事实为 $\text{support}\{\phi(\cdot)\} \subset (-(L-1), 0]$（这里我们用"$\subset$"而不用"$=$"仅仅是因为 $\phi(\cdot)$ 在 $t=0$ 时可以为零也可以不为零）．这两尺度差分方程利用递推的方法可以用来计算尺度函数 $\phi(\cdot)$（11.5 节已详细阐述）．对尺度差分方程两端作傅里叶变换后得到 $\Phi(\cdot)$ 与 $\bar{G}(\cdot)$ 之间的关系，即

$$\Phi(f) = \Phi\left(\frac{f}{2^n}\right)\prod_{m=1}^{n}\frac{\bar{G}\left(\frac{f}{2^m}\right)}{\sqrt{2}}, \quad n = 1, 2, \cdots,$$

其中 $\Phi(\cdot)$ 为 $\phi(\cdot)$ 的傅里叶变换．而 $\bar{G}(\cdot)$ 是 $\{\bar{g}_l\}$ 的傅里叶变换（当后者是一脉冲响应序列时，$\bar{G}(\cdot)$ 变成它的传递函数）．对尺度差分方程两边积分（与假设 $\phi(\cdot)$ 积分后非零结合）给出规范化

$$\bar{G}(0) = \sum_l \bar{g}_l = \sqrt{2}.$$

$\{\phi_{0,k}(\cdot) : k \in \mathbf{Z}\}$ 的规范正交性意味着 $\Phi(\cdot)$ 必须满足式(467a)，即

$$\sum_{m=-\infty}^{\infty} |\Phi(f+m)|^2 = 1.$$

现在，给出如下三个有用的结论：第一，我们发现 $\phi(\cdot)$ 必须积分后为 1 或 -1．选前者后得出结论 $\Phi(0)=1$，从而

$$\Phi(f) = \prod_{m=1}^{\infty}\frac{\bar{G}\left(\frac{f}{2^m}\right)}{\sqrt{2}}.$$

第二，我们有 $|\bar{G}(f)|^2 + \left|\bar{G}\left(f+\frac{1}{2}\right)\right|^2 = 2$（就是式(467b)），所以 $\bar{G}\left(\frac{1}{2}\right) = 0$．第三，对大的 j，第 j 层尺度滤波器的时间逆将有与 $\phi(\cdot)$ 相近的形状，即

$$2^j \bar{g}_{j,-l} = 2^{j/2} g_{j,-l} \approx \phi\left(\frac{l}{2^j}\right).$$

（该逼近由图 471 给出．）

令 $W_0 \subset V_{-1}$ 为 V_0 在 V_{-1} 中的正交补，即如果 $\psi(\cdot) \in W_0$，则 $\psi(\cdot) \in V_{-1}$，但 $\psi(\cdot)$ 与 V_0 中的所有函数正交，即对任意的 $x(\cdot) \in V_0$，有 $\langle \psi(\cdot), x(\cdot) \rangle = 0$．空间 W_0 称为细节空间．如果 $\psi(\cdot) \in L^2(\mathbf{R})$ 使得它的整数平移 $\{\psi_{0,k}(\cdot) : k \in \mathbf{Z}\}$ 形成空间 W_0 的一个规范正交基，则 $\psi(\cdot)$ 称为小波函数．更一般地，子空间 $W_j \subset V_{j-1}$ 是 V_j 在 V_{j-1} 中的正交补，它被认为是相对于尺度 $\tau_j = 2^{j-1}$ 的细节空间．因此

$$V_j = V_{j+1} \oplus W_{j+1}, \quad V_j = W_{j+1} \oplus W_{j+2} \oplus \cdots.$$

信号 $x(\cdot) \in V_0 = V_1 \oplus W_1$ 可以分解到 V_1 与 W_1 的分量为

$$x(t) = s_1(t) + d_1(t),$$

其中

$$d_1(t)=\sum_{k=-\infty}^{\infty}w_{1,k}\psi_{1,k}(t),\quad w_{1,k}\equiv\langle x(\cdot),\psi_{1,k}(\cdot)\rangle,$$

这里 $d_1(\cdot)\in W_1$ 是 $x(\cdot)$ 由粗糙逼近 $s_1(\cdot)$ 失掉的误差细节. $w_{1,k}$ 是单元尺度下的小波系数. $v_{0,k}$ 与 $w_{1,k}$ 之间的关系由下式给出:

$$w_{1,k}=\sum_{m=-\infty}^{\infty}\bar{h}_{m-2k}v_{0,m}=\sum_{l=-\infty}^{\infty}h_l v_{0,2k-l},$$

其中 $\bar{h}_l\equiv\langle\phi_{0,l}(\cdot),\ \psi_{1,0}(\cdot)\rangle$, 且 $h_l=\bar{h}_{-l}$(滤波器 $\{h_l\}$ 是 4.2 节中介绍的小波滤波器, 所以 $\{\bar{h}_l\}$ 是它的时间逆). 我们可以分解 $s_{j-1}(\cdot)\in V_{j-1}=V_j\oplus W_j$ 到在 V_j 与 W_j 中的分量, 得到

$$s_{j-1}(t)=s_j(t)+d_j(t),d_j(t)\equiv\sum_{k=-\infty}^{\infty}w_{j,k}\psi_{j,k}(t),$$

其中 $w_{j,k}\equiv\langle s_{j-1}(\cdot),\ \psi_{j,k}(\cdot)\rangle$. 系数 $w_{j,k}$ 可以用式(480d)递推计算, 即

$$w_{j,k}=\sum_{m=-\infty}^{\infty}\bar{h}_{m-2k}v_{j-1,m}=\sum_{l=-\infty}^{\infty}h_l v_{j-1,2k-l}.$$

因为

$$V_{j_0}=V_j\oplus W_j\oplus\cdots\oplus W_{j_0+1},$$

我们知道, 对某个 $j_0\in\mathbf{Z}$ 如果 $x(\cdot)\in V_{j_0}$, 则对任一 $j>j_0$,

$$x(t)=\sum_{k=-\infty}^{\infty}v_{j,k}\phi_{j,k}(t)+\sum_{l=j_0+1}^{j}\sum_{k=-\infty}^{\infty}w_{l,k}\psi_{l,k}(t)$$

和

$$\int_{-\infty}^{\infty}x^2(t)\mathrm{d}t=\sum_{k=-\infty}^{\infty}v_{j,k}^2+\sum_{l=j_0+1}^{j}\sum_{k=-\infty}^{\infty}w_{l,k}^2.$$

$j-1$ 层尺度系数可以通过 j 层尺度系数和小波系数使用下式重构:

$$v_{j-1,k}=\sum_{m=-\infty}^{\infty}\bar{g}_{k-2m}v_{j,m}+\sum_{m=-\infty}^{\infty}\bar{h}_{k-2m}w_{j,m}.$$

通过式(472), 小波函数 $\psi(\cdot)$ 与尺度函数 $\phi(\cdot)$ 有如下关系:

$$\psi(t)\equiv\sqrt{2}\sum_{l=-\infty}^{\infty}\bar{h}_l\phi(2t-l),\text{其中 }\bar{h}_l\equiv(-1)^l\bar{g}_{1-l-L}.$$

以上结论类似于两尺度差分方程．小波滤波器$\{\bar{h}_l\}$，在 $-(L-1)\leqslant l\leqslant 0$ 之外为零，是相应于尺度滤波器$\{\bar{g}_l\}$，在 $-(L-1)\leqslant l\leqslant 0$ 之外也为零的正交镜像滤波器(QMF)．这正交镜像滤波器的定义确保了 $\bar{h}_{-l}=h_l$．上面 $\psi(\cdot)$借助于$\phi_{-1,l}(\cdot)$的表达式以及 support$\{\phi(\cdot)\}\subset(-(L-1),\ 0]$ 意味着对 $\psi(\cdot)$的支撑也是同样的情况，即 support$\{\psi(\cdot)\}\subset(-(L-1),\ 0]$．通过 11.7 节中定义的递推方法，上述表达式可以用来计算小波函数 $\psi(\cdot)$的值．$\psi(\cdot)$的傅里叶变换 $\Psi(\cdot)$如下：

$$\Psi(f)=\frac{\bar{H}\left(\frac{f}{2}\right)}{\sqrt{2}}\prod_{m=2}^{\infty}\frac{\bar{G}\left(\frac{f}{2^m}\right)}{\sqrt{2}},$$

其中 $\bar{H}(\cdot)$是$\{\bar{h}_l\}$的傅里叶变换．正交镜像滤波器关系推出

$$\bar{H}(f)=-\,\mathrm{e}^{\mathrm{i}2\pi f(L-1)}\bar{G}\left(\frac{1}{2}-f\right),$$

并且，由于 $\bar{G}\left(\frac{1}{2}\right)=0$，由此得出 $\bar{H}(0)=0$，因此 $\Psi(0)=\int_{-\infty}^{\infty}\psi(t)\mathrm{d}t=0$．对大的 j，第 j 层小波滤波器的时间逆与 $\psi(\cdot)$有相似的形状，即

$$2^j\bar{h}_{j,-l}=2^{j/2}h_{j,-l}\approx\psi\left(\frac{l}{2^j}\right).$$

(这个逼近在图 478 中展示．)

如果 $\psi(\cdot)$有 r 阶消失距，即

$$\int_{-\infty}^{\infty}t^m\psi(t)\mathrm{d}t=0,\text{则}\ \bar{H}^{(m)}(0)=0,\quad \sum_{l=-\infty}^{\infty}l^m\bar{h}_l=0$$

对 $m=0,\ 1,\ \cdots,\ r-1$ 成立．因此滤波器序列$\{\bar{h}_l\}$也可以称为有 r 阶离散消失距．进而，对于 $m=0,\ 1,\ \cdots,\ r-1$，$\bar{H}^{(m)}(0)=0$ 意味着对 $m=0,\ 1,\ \cdots,\ r-1$ 也有 $\bar{G}^{(m)}\left(\frac{1}{2}\right)=0$，因此$\{\bar{g}_l\}$的 z 变换$\bar{\bar{G}}(z)$有 r 个根为 -1．当$|z|=1$ 时，即 $z=\mathrm{e}^{\mathrm{i}2\pi f}$，$z$ 变换的模的平方有如下形式：

$$|\bar{\bar{G}}(z)|^2=\bar{\bar{G}}(z)\bar{\bar{G}}(z^{-1})=g_{-(L-1)}^2(z^{-1}+2+z)^r\prod_{l=r+1}^{L-1}(z-z_l)(z^{-1}-z_l).$$

如果令$\mathcal{G}(f)\equiv|\bar{G}(f)|^2$，其中 $\bar{G}(f)=\bar{\bar{G}}(\mathrm{e}^{\mathrm{i}2\pi f})$，则相应的滤波器系数$\{\bar{g}_l\}$的平方增益函数由下式给出：

$$\mathcal{G}(f)=\left|\frac{1+\mathrm{e}^{\mathrm{i}2\pi f}}{2}\right|^{2r}|Q(f)|^2=\cos^{2r}(\pi f)\,\mathcal{Q}(f),$$

其中$\mathcal{Q}(f)$是 $\cos(2\pi f)$的多项式．Daubechies(1988)构造了对于 $L=2r$ 的一组尺度滤波器，系数由下式的平方增益函数给出：

$$\mathcal{G}(f)=2\cos^{2r}(\pi f)\sum_{k=0}^{r-1}\binom{r-1+k}{k}\sin^{2k}(\pi f).$$

表 499

$h_l \equiv \tilde{h}_{-l}$，$l=0$，…，$L-1$	$g_l \equiv \tilde{g}_{-l}$，$l=0$，…，$L-1$
$\tilde{h}_l = (-1)^l \tilde{g}_{1-l-L}$	$\tilde{g}_l \equiv (-1)^{l+1} \tilde{h}_{1-l-L}$
$\{\tilde{h}_l\} \leftrightarrow \tilde{H}(\cdot)$	$\{\tilde{g}_l\} \leftrightarrow \tilde{G}(\cdot)$
$\{h_l\} \leftrightarrow H(\cdot)$	$\{g_l\} \leftrightarrow G(\cdot)$
$H(f) = \tilde{H}(-f)$	$G(f) = \tilde{G}(-f)$
$\tilde{H}(0) = 0$	$\tilde{G}(0) = \sqrt{2}$
$\tilde{H}(f) = -\mathrm{e}^{i2\pi f(L-1)} \tilde{G}\left(\frac{1}{2} - f\right)$	$\tilde{G}(f) = \mathrm{e}^{i2\pi f(L-1)} \tilde{H}\left(\frac{1}{2} - f\right)$
$\tilde{H}^{(m)}(0) = 0 \quad m=0, \cdots, r-1$	$\tilde{G}^{(m)}\left(\frac{1}{2}\right) = 0 \quad m=0, \cdots, r-1$
$\tilde{h}_l = \int \phi(t-l) \frac{\psi\left(\frac{t}{2}\right)}{\sqrt{2}} \mathrm{d}t$	$\tilde{g}_l = \int \phi(t-l) \frac{\phi\left(\frac{t}{2}\right)}{\sqrt{2}} \mathrm{d}t$
$\int \psi(t) \mathrm{d}t = 0$	$\int \phi(t) \mathrm{d}t = 1$
support$\{\psi(\cdot)\} \subset (-(L-1), 0]$	support$\{\phi(\cdot)\} \subset (-(L-1), 0]$
$\psi(\cdot) \leftrightarrow \Psi(\cdot)$	$\phi(\cdot) \leftrightarrow \Phi(\cdot)$
$\Psi(-2^j f) \approx \tilde{H}_j(f)$	$\Phi(-2^j f) \approx \tilde{G}_j(f)$
$\psi\left(-\frac{l}{2^j}\right) \approx 2^j \tilde{h}_{j,l} = 2^{j/2} h_{j,l}$	$\phi\left(-\frac{l}{2^j}\right) \approx 2^j \tilde{g}_{j,l} = 2^{j/2} g_{j,l}$
$\Psi(f) = \Phi\left(\frac{f}{2}\right) \frac{\tilde{H}\left(\frac{f}{2}\right)}{\sqrt{2}}$	$\Phi(f) = \Phi\left(\frac{f}{2}\right) \frac{\tilde{G}\left(\frac{f}{2}\right)}{\sqrt{2}}$
$\Psi(f) = \frac{\tilde{H}\left(\frac{f}{2}\right)}{\sqrt{2}} \prod_{m=2}^{\infty} \frac{G\left(\frac{f}{2^m}\right)}{\sqrt{2}}$	$\Phi(f) = \prod_{m=1}^{\infty} \frac{\tilde{G}\left(\frac{f}{2^m}\right)}{\sqrt{2}}$
$\psi_{j,k}(t) \equiv \psi\left(\frac{t}{2^j} - k\right) \Big/ \sqrt{2^j}$	$\phi_{j,k}(t) \equiv \phi\left(\frac{t}{2^j} - k\right) \Big/ \sqrt{2^j}$
$\psi(t) = \sqrt{2} \sum_l \tilde{h}_l \phi(2t-l)$	$\phi(t) = \sqrt{2} \sum_l \tilde{g}_l \phi(2t-l)$
$\omega_{j,k} = \int x(t) \psi_{j,k}(t) \mathrm{d}t$	$v_{j,k} = \int x(t) \phi_{j,k}(t) \mathrm{d}t$
$\omega_{j,k} = \sum_l h_l v_{j-1,2k-l}$	$v_{j,k} = \sum_l g_l v_{j-1,2k-l}$

注：(i)小波与尺度滤波器$\{h_l\}$和$\{g_l\}$以及它们的时间逆$\{\tilde{h}_l\}$和$\{\tilde{g}_l\}$；(ii)小波函数 $\psi(\cdot)$和尺度函数 $\phi(\cdot)$的关键关系．

对于一个给定的 r，这个表达式可以再写成如下形式：

$$|\bar{G}(z)|^2 = \bar{g}^2_{-(L-1)}|z+1|^L\bar{Q}_0(z)\bar{Q}_0(z^{-1}),$$

其中 $\bar{Q}_0(z)$ 是 z 的 $\frac{L}{2}-1$ 次非负指数实系数多项式，并且可以写成

$$\bar{Q}_0(z) = \prod_{l=1}^{K_0}(z-r_l)\prod_{l=1}^{K_1}(z-z_l)(z-z_l^*);$$

这里 r_l 和 z_l 分别是实根和复根，且 $K_0+2K_1=\frac{L}{2}-1$. $\{\bar{g}_l\}$ 的 z 变换 $\bar{G}(z)$ 用"谱分解"构造，这里我们选择 $r=\frac{L}{2}$ 个根为 -1，之后使用 $\frac{L}{2}-1$ 个在单位圆外(相应于"最大延时"偶极子)的根来确定 $\bar{Q}_0(z)$ 的一个分解 $\bar{Q}_0(z)\bar{Q}_0(z^{-1})$. 得到的滤波器 $\{g_l=\bar{g}_{-l}\}$ 是 D(L)尺度滤波器，且有一个具有 $\frac{L}{2}$ 个根为 -1，$\frac{L}{2}-1$ 个根在单位圆内(相应于"最小延时"偶)的 z 变换. 这一方法可以通过修改来产生 LA(L)尺度滤波器：我们选择最大和最小延时偶极子的联合而不是选择 $\bar{G}(z)$ 的 $\frac{L}{2}-1$ 个最大延时偶极子，使得到的滤波器 $\{\bar{g}_l\}$ 提升到 $\{g_l\}$ 以便满足式(112a)的最优性判别标准.

最后，注意表 499 总结了涉及小波与尺度滤波器的关键关系、小波和尺度函数以及相关量.

11.12 练习

[11.1] 使用 11.6 节中列出的方法的推广来构造一个与图 471 类似的对于 D(6)的图(D(6)尺度滤波器的系数由表 109 给出).

[11.2] 证明 $\psi_{0,k}(\cdot)(k\in\mathbf{Z})$ 构成 W_0 的一个规范正交基. 提示：给定 $\int_{-\infty}^{\infty}\psi_{0,k}(t)\phi_{0,l}(t)\mathrm{d}t=0$ 和练习[474b]的结果，需要证明任一 $x(\cdot)\in V_{-1}\subset L^2(\mathbf{R})$ 可以写成 $x(t)=\sum_k c_k^\phi\phi_{0,k}(t)+\sum_k c_k^\psi\psi_{0,k}(t)$，其中系数 $\{c_k^\phi\}$ 和 $\{c_k^\psi\}$ 必须满足 $\sum_k(c_k^\phi)^2<\infty$ 和 $\sum_k(c_k^\psi)^2<\infty$. 这个使用哈尔尺度和小波函数的特殊情况很容易证明.

[11.3] 验证式(477c).

[11.4] 假设 $x(\cdot)\in L^2(\mathbf{R})$ 和 V_j，$j\in\mathbf{Z}$，是一个具有相应尺度函数 $\phi(\cdot)$ 的多分辨分析. 考虑由下式给出的关于 $x(\cdot)$ 在子空间 V_j 上的投影的逼近：

$$s_j(t)=\sum_{k=-\infty}^{\infty}v_{j,k}\phi_{j,k}(t),\quad 其中\ v_{j,k}\equiv\langle x(\cdot),\phi_{j,k}(\cdot)\rangle.$$

对任一 $j_0<j$，并且证明对这个 $x(\cdot)$，式(481d)的右端能够被认为是 $s_{j_0}(t)$，即 $x(\cdot)$ 在子空间 V_{j_0} 上的投影.

[11.5] 验证式(493a)和(493b)的形式.

附录

嵌入练习答案

在这里，我们给出嵌入各章正文的所有习题的解答(想给出各章末尾的练习的解答的教师请与我们联系). 我们认为，读者通过努力做这些练习可以帮助理解小波，所以我们鼓励读者在看解答之前先做完练习. 通过做练习读者有机会找到更优的解答，并且我们非常乐意听取大家关于错误或改进的宝贵意见(e-mail 和通信地址在书的前面已经给出).

练习[22a]答案 首先注意，由于 t 和 k 都是整数，$e^{-i2\pi kt}=\cos(2\pi kt)-i\sin(2\pi kt)=1$. 因此，

$$
\begin{aligned}
A(f+k) &= \sum_{t=-\infty}^{\infty} a_t e^{-i2\pi(f+k)t} \\
&= \sum_{t=-\infty}^{\infty} a_t e^{-i2\pi ft} e^{-i2\pi kt} \\
&= \sum_{t=-\infty}^{\infty} a_t e^{-i2\pi ft} = A(f).
\end{aligned}
$$

练习[22b]答案 一方面，我们有

$$
\begin{aligned}
A(-f) &= \sum_{t=-\infty}^{\infty} a_t e^{-i2\pi(-f)t} \\
&= \sum_{t=-\infty}^{\infty} a_t e^{i2\pi ft};
\end{aligned}
$$

另一方面，因为对于实值变量有 $a_t^*=a_t$，

$$
\begin{aligned}
A^*(f) &= \Big(\sum_{t=-\infty}^{\infty} a_t e^{-i2\pi ft}\Big)^* \\
&= \sum_{t=-\infty}^{\infty} a_t^* (e^{-i2\pi ft})^* \\
&= \sum_{t=-\infty}^{\infty} a_t e^{i2\pi ft}.
\end{aligned}
$$

练习[22c]答案 利用 $A(f)$的定义，我们有

$$
\begin{aligned}
\int_{-1/2}^{1/2} A(f) e^{i2\pi ft}\,df &= \int_{-1/2}^{1/2} \Big(\sum_{t'=-\infty}^{\infty} a_{t'} e^{-i2\pi ft'}\Big) e^{i2\pi ft}\,df \\
&= \sum_{t'=-\infty}^{\infty} a_{t'} \int_{-1/2}^{1/2} e^{i2\pi f(t-t')}\,df.
\end{aligned}
$$

如果 $t=t'$，则上面的积分为 1；另一方面，因为 $t-t'$ 也是整数，那么

$$\int_{-1/2}^{1/2} e^{i2\pi f(t-t')}\,df = \frac{1}{i2\pi(t-t')}e^{i2\pi f(t-t')}\Big|_{-1/2}^{1/2}$$

$$= \frac{1}{i2\pi(t-t')}(e^{i\pi(t-t')} - e^{-i\pi(t-t')})$$

$$= \frac{\sin(\pi(t-t'))}{\pi(t-t')} = 0,$$

因为上面无限和中所有项除了 $t=t'$ 外均为 0，在 $t=t'$ 时为 a_t.

练习[23a]答案 因为

$$A(f) = \sum_{t=-\infty}^{\infty} a_t e^{-i2\pi ft},$$

$$B^*(f) = \sum_{t'=-\infty}^{\infty} b_{t'}^* e^{i2\pi ft'},$$

我们有

$$\int_{-1/2}^{1/2} A(f)B^*(f)df = \int_{-1/2}^{1/2}\Big(\sum_{t=-\infty}^{\infty} a_t e^{-i2\pi ft}\Big)\Big(\sum_{t'=-\infty}^{\infty} b_{t'}^* e^{i2\pi ft'}\Big)df$$

$$= \sum_{t=-\infty}^{\infty} a_t \sum_{t'=-\infty}^{\infty} b_{t'}^* \int_{-1/2}^{1/2} e^{i2\pi f(t'-t)}\,df.$$

用与练习[22c]相同的论证，当 $t=t'$ 时，上面的积分为 1，其他情况为 0，因此里面的和正好减化为 b_t^*.

练习[23b]答案 因为

$$a_t = \int_{-1/2}^{1/2} A(f)e^{i2\pi ft}\,df,$$

我们有(令 $f'=2f$)

$$a_{2n} = \int_{-1/2}^{1/2} A(f)e^{i2\pi f(2n)}\,df$$

$$= \frac{1}{2}\int_{-1}^{1} A\Big(\frac{f'}{2}\Big)e^{i2\pi f'n}\,df'$$

$$= \frac{1}{2}\Big[\int_{-1}^{-1/2} + \int_{-1/2}^{1/2} + \int_{1/2}^{1} A\Big(\frac{f'}{2}\Big)e^{i2\pi f'n}\,df'\Big].$$

在括号中的第一个和第三个积分，分别改变积分变量 $f=f'+1$ 和 $f=f'-1$，然后组合结果得到

$$a_{2n} = \int_{-1/2}^{1/2} \frac{1}{2}\Big[A\Big(\frac{f}{2}\Big)+A\Big(\frac{f}{2}+\frac{1}{2}\Big)\Big]e^{i2\pi fn}\,df$$

(因为单位周期性，并且 $\exp(\pm i2\pi n)=1$，其中 n 为整数，我们还可以利用事实 $A\Big(\frac{f}{2}-\frac{1}{2}\Big) = A\Big(\frac{f}{2}+\frac{1}{2}\Big)$). 上面的积分具有离散傅里叶逆变换的形式，它告诉我们 $\{a_{2n}\}$ 的离散傅里叶变换像在练习中阐述的.

练习[24]答案 使用$\{a * b_t\}$的定义，像需要的，我们有(令 $v=t-u$)

$$\begin{aligned}\sum_{t=-\infty}^{\infty} a * b_t \mathrm{e}^{-\mathrm{i}2\pi ft} &= \sum_{t=-\infty}^{\infty}\Big(\sum_{u=-\infty}^{\infty} a_u b_{t-u}\Big)\mathrm{e}^{-\mathrm{i}2\pi ft}\\ &= \sum_{u=-\infty}^{\infty} a_u\Big(\sum_{t=-\infty}^{\infty} b_{t-u}\mathrm{e}^{-\mathrm{i}2\pi ft}\Big)\\ &= \sum_{u=-\infty}^{\infty} a_u\Big(\sum_{v=-\infty}^{\infty} b_v \mathrm{e}^{-\mathrm{i}2\pi f(v+u)}\Big)\\ &= \sum_{u=-\infty}^{\infty} a_u \mathrm{e}^{-\mathrm{i}2\pi fu}\Big(\sum_{v=-\infty}^{\infty} b_v \mathrm{e}^{-\mathrm{i}2\pi fv}\Big)\\ &= A(f)B(f).\end{aligned}$$

练习[27]答案 令 $B(\cdot)$是级联滤波器的输入的傅里叶变换．从练习[24]得到，第一个滤波器的输出具有由 $B(f)A_1(f)$定义的离散傅里叶变换．由于输入的第二个滤波器的离散傅里叶变换还由 $B(f)A_1(f)$定义，再次利用练习[24]可知，第二个滤波器的输出的离散傅里叶变换用 $B(f)A_1(f)A_2(f)$定义．重复 M 次，就可以得到级联滤波器的输出$\{c_t\}$的离散傅里叶变换为 $C(f)=B(f)A_1(f)A_2(f)\cdots A_M(f)$.

令

$$A(f) \equiv \prod_{m=1}^{M} A_m(f),$$

则有 $C(f)=B(f)A(f)$．令$\{a_t\}$是 $A(\cdot)$的傅里叶逆变换．从练习[24]可知$\{c_t\}$是$\{b_t\}$和$\{a_t\}$的卷积．从练习[24]和归纳可知$\{a_t\}$是$\{a_{m,t}\}$，$m=1$，…，M 的卷积．

练习[28a]答案 首先，我们证明 $M=2$ 时的结果成立，从练习[27]可知级联滤波器的传递函数为

$$\begin{aligned}A(f) &= A_1(f)A_2(f)\\ &= \Big(\sum_{t=K_1}^{K_1+L_1-1} a_{1,t}\mathrm{e}^{-\mathrm{i}2\pi ft}\Big)\Big(\sum_{u=K_2}^{K_2+L_2-1} a_{2,u}\mathrm{e}^{-\mathrm{i}2\pi fu}\Big)\\ &= \sum_{t=K_1}^{K_1+L_1-1}\ \sum_{u=K_2}^{K_2+L_2-1} a_{1,t}a_{2,u}\mathrm{e}^{-\mathrm{i}2\pi f(t+u)}.\end{aligned}$$

现在，当 $t=K_1+L_1-1$，$u=K_2+L_2-1$ 时得到 $t+u$ 的最大值；同时，当 $t=K_1$，$u=K_2$ 时得到 $t+u$ 的最小值．令$\{a_t\}$是 $A(\cdot)$的离散傅里叶逆变换，从而得到

$$A(f) = \sum_{v=K_1+K_2}^{K_1+L_1-1+K_2+L_2-1} a_v \mathrm{e}^{-\mathrm{i}2\pi fv},$$

所以$\{a_t\}$的宽度为

$$K_1+L_1-1+K_2+L_2-1-(K_1+K_2)+1 = L_1+L_2-1,$$

这就建立了 $M=2$ 的情况.

其次讨论，如果对于 $M-1$ 个滤波器结果成立，那么对于 M 个滤波器必定成立. 从练习[27]可知，由 M 个滤波器组成的级联滤波器的等价滤波器的传递函数为

$$A(f)=\prod_{m=1}^{M}A_m(f)=A_M(f)\prod_{m=1}^{M-1}A_m(f)\equiv A_M(f)\overline{A}(f),$$

其中 $\overline{A}(\cdot)$由 $M-1$ 个滤波器组成的级联滤波器的等价滤波器的传递函数，这 $M-1$ 个滤波器对应的传递函数分别为 $A_1(\cdot)$，…，$A_{M-1}(\cdot)$. 由归纳假设可知$\{\overline{a}_t\}\leftrightarrow\overline{A}(\cdot)$宽为 $L'=\sum_{m=1}^{M-1}L_m-M+2$. 因为 $A_M(f)\overline{A}(f)$定义了 2 个滤波器的级联滤波器的传递函数，所以$M=2$ 的情况说明$\{a_t\}\leftrightarrow A(\cdot)$的宽为

$$\begin{aligned}L'+L_M-1&=\sum_{m=1}^{M-1}L_m-M+2+L_M-1\\&=\sum_{m=1}^{M}L_m-M+1.\end{aligned}$$

练习[28b]答案 因为对于所有整数 m 有 $\exp(i2\pi m)=1$，并且因为乘积 tn 为整数，像需要的，我们有

$$A_{k+nN}=\sum_{t=0}^{N-1}a_t e^{-i2\pi t(k+nN)/N}=\sum_{t=0}^{N-1}a_t e^{-i2\pi tk/N}e^{-i2\pi tn}=\sum_{t=0}^{N-1}a_t e^{-i2\pi tk/N}=A_k.$$

练习[29a]答案 为了证明提示，注意

$$(1-z)\sum_{u=0}^{N-1}z^u=\sum_{u=0}^{N-1}z^u-\sum_{u=0}^{N-1}z^{u+1}=1-z^N,$$

由此能够推出，如果 $z\neq1$，则

$$\sum_{u=0}^{N-1}z^u=\frac{1-z^N}{1-z}.$$

另一方面，如果 $z=1$，则 $\sum_{u=0}^{N-1}z^u=N$. 事实上，如果 $z=e^{i\theta}$，则 $z=1$ 当且仅当 $\theta=0$，$\pm2\pi$，$\pm4\pi$，…. 利用 A_k 的定义，我们有

$$\begin{aligned}\frac{1}{N}\sum_{k=0}^{N-1}A_k e^{i2\pi tk/N}&=\frac{1}{N}\sum_{k=0}^{N-1}\Big(\sum_{u=0}^{N-1}a_u e^{-i2\pi uk/N}\Big)e^{i2\pi tk/N}\\&=\frac{1}{N}\sum_{u=0}^{N-1}a_u\sum_{k=0}^{N-1}e^{i2\pi(t-u)k/N}.\end{aligned}$$

如果 $t=u$，里面的和为 N；如果 $t\neq u$，注意 $z\equiv\exp(i2\pi(t-u)/N)\neq1$(由于 $0<2\pi|t-u|/N\leqslant2\pi(N-1)/N<2\pi$)，因此，里面的和为

$$\sum_{k=0}^{N-1}e^{i2\pi(t-u)k/N}=\frac{1-e^{i2\pi(t-u)}}{1-e^{i2\pi(t-u)}/N}=0,$$

由于 $\exp(i2\pi(t-u))=1$，这是因为 $t-u$ 始终为整数. 因此二重和正好为 a_t，从而得到所要求的结果.

练习[29b]答案 因为

$$A_k = \sum_{t=0}^{N-1} a_t e^{-i2\pi tk/N}$$

$$B_k^* = \sum_{u=0}^{N-1} b_u^* e^{i2\pi uk/N},$$

我们有

$$\frac{1}{N}\sum_{k=0}^{N-1} A_k B_k^* = \frac{1}{N}\sum_{k=0}^{N-1}\left(\sum_{t=0}^{N-1} a_t e^{-i2\pi tk/N}\right)\left(\sum_{u=0}^{N-1} b_u^* e^{i2\pi uk/N}\right)$$

$$= \frac{1}{N}\sum_{t=0}^{N-1} a_t \sum_{u=0}^{N-1} b_u^* \sum_{k=0}^{N-1} e^{i2\pi(u-t)k/N}.$$

在练习[29a]中，我们建立了

$$\sum_{k=0}^{N-1} e^{i2\pi(u-t)k/N} = \begin{cases} N, & u=t; \\ 0, & \text{其他}. \end{cases}$$

因此三重和化简为一个单重和，从而得到结果.

练习[30]答案 $a * b_t$ 的定义产生

$$\sum_{t=0}^{N-1} a * b_t e^{-i2\pi tk/N} = \sum_{t=0}^{N-1}\left(\sum_{u=0}^{N-1} a_u b_{t-u \bmod N}\right) e^{-i2\pi tk/N}$$

$$= \sum_{u=0}^{N-1} a_u \sum_{t=0}^{N-1} b_{t-u \bmod N} e^{-i2\pi tk/N}.$$

现在，无限序列$\{b_{t-u \bmod N} e^{-i2\pi tk/N};\ t=\cdots, -1, 0, 1, \cdots\}$是以 N 为周期的周期序列，因此在序列中在任何 N 个邻接的变量上和是相同的. 另外利用事实 $b_{t \bmod N} = b_t$，$0 \leqslant t \leqslant N-1$，如需要的，我们有

$$\sum_{t=0}^{N-1} a * b_t e^{-i2\pi tk/N} = \sum_{u=0}^{N-1} a_u \sum_{t=0}^{N-1} b_t e^{-i2\pi(t+u)k/N}$$

$$= \sum_{u=0}^{N-1} a_u e^{-i2\pi uk/N} \sum_{t=0}^{N-1} b_t e^{-i2\pi tk/N} = A_k B_k.$$

练习[31]答案 令$\tilde{a}_0 = a_0^*$ 和$\tilde{a}_t = a_{N-t}^*$，对于 $t=1, \cdots, N-1$，作为技巧. 这个可以从图 31a 和 31b 中用圆圈标示的元素对照的元素可以看到，或者从下面形式的论证可以看出来. 令 $v=-u$，我们有

$$a^* \star b_t \equiv \sum_{u=0}^{N-1} a_u^* b_{u+t \bmod N}$$

$$= \sum_{v=-(N-1)}^{0} a_{-v}^* b_{t-v \bmod N}$$

$$= \sum_{v=1}^{N} a_{N-v}^* b_{t-v \bmod N} \qquad (\text{因为 } b_{t-v+N \bmod N} = b_{t-v \bmod N})$$

$$= \sum_{v=0}^{N-1} a_{N-v \bmod N}^* b_{t-v \bmod N} \equiv \sum_{v=0}^{N-1} \tilde{a}_v b_{t-v \bmod N},$$

其中$\tilde{a}_t = a^*_{N-t \bmod N}$(回忆$\tilde{a}_0 = a^*_{N \bmod N} = a^*_0$，而$\tilde{a}_t = a^*_{N-t}$，$t=1, \cdots, N-1$)．最后，$\{\tilde{a}_t\}$的傅里叶变换由

$$\begin{aligned}\sum_{t=0}^{N-1} \tilde{a}_t e^{-i2\pi tk/N} &= a_0^* + \sum_{t=1}^{N-1} a^*_{N-t} e^{-i2\pi tk/N} \\ &= a_0^* + \sum_{u=1}^{N-1} a_u^* e^{-i2\pi(N-u)k/N} \qquad (\text{令 } u = N-t) \\ &= \sum_{u=0}^{N-1} a_u^* e^{i2\pi uk/N} = \Big(\sum_{u=0}^{N-1} a_u e^{-i2\pi uk/N}\Big)^* = A_k^*\end{aligned}$$

给出，如所需要的．

练习[33]答案 对于所有整数n，k，使用事实

$$e^{-i2\pi(u+nN)k/N} = e^{-i2\pi uk/N} e^{-i2\pi nk} = e^{-i2\pi uk/N},$$

我们有

$$\begin{aligned}A_k^\circ \equiv \sum_{u=0}^{N-1} a_u^\circ e^{-i2\pi uk/N} &= \sum_{u=0}^{N-1} \Big(\sum_{n=-\infty}^{\infty} a_{u+nN}\Big) e^{-i2\pi uk/N} \\ &= \sum_{n=-\infty}^{\infty} \sum_{u=0}^{N-1} a_{u+nN} e^{-i2\pi(u+nN)k/N} \\ &= \sum_{t=-\infty}^{\infty} a_t e^{-i2\pi tk/N} = A\Big(\frac{k}{N}\Big),\end{aligned}$$

在此，二重和化简为单重和是由于在两个表达式中无穷序列$\{a_t e^{-i2\pi tk/N}\}$的每项恰好出现一次(二重和增加了N个邻接的变量组成的组，然后把这所有的组再加在一起)．

练习[34]答案 作为一个简单的反例，考虑下面的三个滤波器，对每个滤波器而言，当$t=-1, -2, \cdots$，或者$t=4, 5, \cdots$时，$a_t=0$：

$$a_t = \begin{cases} \dfrac{1}{2}, & t = 0,1; \\ -\dfrac{1}{2}, & t = 2,3; \end{cases} \qquad a_t = \begin{cases} \dfrac{1-\sqrt{3}}{4\sqrt{2}}, & t = 0; \\ \dfrac{-3+\sqrt{3}}{4\sqrt{2}}, & t = 1; \\ \dfrac{3+\sqrt{3}}{4\sqrt{2}}, & t = 2; \\ \dfrac{-1-\sqrt{3}}{4\sqrt{2}}, & t = 3; \end{cases}$$

和

$$a_t = 1/2, \quad 0 \leqslant t \leqslant 3$$

(这些滤波器是第4章所讨论的第1层和第2层小波和尺度滤波器的特殊情况)．上面每个滤波器满足$\sum a_t^2 = 1$．对上面以长为$N=2$进行周期化产生滤波器$\{a_0^\circ = 0, a_1^\circ = 0\}$，$\{a_0^\circ = 1/\sqrt{2}, a_1^\circ = -1/\sqrt{2}\}$和$\{a_0^\circ = 1, a_1^\circ = 1\}$，对此，我们分别有

$$\sum_{t=0}^{1} |a_t^\circ|^2 = 0; \qquad \sum_{t=0}^{1} |a_t^\circ|^2 = 1; \qquad \sum_{t=0}^{1} |a_t^\circ|^2 = 2.$$

这个例子说明了，周期化的滤波器的平方和小于，等于或者大于原来的滤波器的平方和.

这里是更正式的证明. 因为

$$\{a_t^\circ: t=0,\cdots,N-1\}\leftrightarrow\left\{A\left(\frac{k}{N}\right): k=0,\cdots,N-1\right\},$$

帕塞瓦尔定理表明

$$\sum_{t=0}^{N-1}|a_t^\circ|^2=\frac{1}{N}\sum_{k=0}^{N-1}\left|A\left(\frac{k}{N}\right)\right|^2$$

(见式(36h)). 因为 $A(\cdot)$ 是非周期化滤波器 $\{a_t\}$ 的离散傅里叶变换，所以可以写

$$\sum_{t=0}^{N-1}|a_t^\circ|^2=\frac{1}{N}\sum_{k=0}^{N-1}\left(\sum_{t=-\infty}^{\infty}a_t e^{-i2\pi tk/N}\right)\left(\sum_{u=-\infty}^{\infty}a_u^* e^{i2\pi uk/N}\right)$$

$$=\frac{1}{N}\sum_{t=-\infty}^{\infty}\sum_{u=-\infty}^{\infty}a_t a_u^*\sum_{k=0}^{N-1}e^{i2\pi(u-t)k/N}.$$

使用类似练习[29a]的解法给出的论证，最里面的和仅当 $u-t$ 是 N 的整数倍时非 0，在这种情况下这个和为 N. 对给定的 t，当 $u=t$，$t\pm N$，$t\pm 2N$，…时就是这种情况，因此，一般地，我们有

$$\sum_{t=0}^{N-1}|a_t^\circ|^2=\sum_{t=-\infty}^{\infty}a_t\sum_{n=-\infty}^{\infty}a_{t+nN}^*$$

$$=\sum_{t=-\infty}^{\infty}|a_t|^2+\sum_{t=-\infty}^{\infty}a_t\sum_{n=1}^{\infty}(a_{t+nN}^*+a_{t-nN}^*)$$

$$\neq\sum_{t=-\infty}^{\infty}|a_t|^2.$$

练习[43]答案 注意到 $\widetilde{\mathcal{O}}\equiv\mathcal{O}^T$ 是正交矩阵，因此式(43c)告诉我们

$$\boldsymbol{X}=\sum_{k=0}^{N-1}\langle\boldsymbol{X},\widetilde{\mathcal{O}}_{k\cdot}\rangle\,\widetilde{\mathcal{O}}_{k\cdot};$$

然而 $\widetilde{\mathcal{O}}_{k\cdot}=\mathcal{O}_{\cdot k}$，所以像要求的，

$$\boldsymbol{X}=\sum_{k=0}^{N-1}\langle\boldsymbol{X},\mathcal{O}_{\cdot k}\rangle\,\mathcal{O}_{\cdot k}.$$

练习[44a]答案 因为当 $j\neq j'$ 时 $\mathcal{O}_{j\cdot}^T\mathcal{O}_{j'\cdot}=0$，像要求的我们有

$$\|\boldsymbol{X}\|^2=\left\|\sum_{j=0}^{N-1}O_j\,\mathcal{O}_{j\cdot}\right\|^2$$

$$=\left(\sum_{j=0}^{N-1}O_j\,\mathcal{O}_{j\cdot}\right)^T\left(\sum_{j'=0}^{N-1}O_{j'}\,\mathcal{O}_{j'\cdot}\right)$$

$$=\sum_{j=0}^{N-1}\sum_{j'=0}^{N-1}O_jO_{j'}\,\mathcal{O}_{j\cdot}^T\mathcal{O}_{j'\cdot}$$

$$=\sum_{j=0}^{N-1}O_j^2\,\mathcal{O}_{j\cdot}^T\mathcal{O}_{j\cdot}$$

$$=\sum_{j=0}^{N-1}(O_j\,\mathcal{O}_{j\cdot})^T(O_j\,\mathcal{O}_{j\cdot})$$

$$=\sum_{j=0}^{N-1}\|O_j\,\mathcal{O}_{j\cdot}\|^2.$$

练习[44b]答案 因为由定义 $e=X-\hat{X}$，我们有

$$\|e\|^2=\|X-\hat{X}\|^2=\left\|\sum_{j=0}^{N-1}O_j\mathcal{O}_{j\cdot}-\sum_{j=0}^{N'-1}\alpha_j\mathcal{O}_{j\cdot}\right\|^2$$

$$=\left\|\sum_{j=0}^{N'-1}(O_j-\alpha_j)\mathcal{O}_{j\cdot}+\sum_{j=N'}^{N-1}O_j\mathcal{O}_{j\cdot}\right\|^2$$

$$=\left\|\sum_{j=0}^{N-1}\beta_j\mathcal{O}_{j\cdot}\right\|^2,$$

其中

$$\beta_j=\begin{cases}O_j-\alpha_j, & j=0,\cdots,N'-1;\\ O_j, & j=N',\cdots,N-1.\end{cases}$$

现在

$$\left\|\sum_{j=0}^{N-1}\beta_j\mathcal{O}_{j\cdot}\right\|^2=\left(\sum_{j=0}^{N-1}\beta_j\mathcal{O}_{j\cdot}\right)^{\mathrm{T}}\left(\sum_{k=0}^{N-1}\beta_k\mathcal{O}_{k\cdot}\right)$$

$$=\sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\beta_j\beta_k\mathcal{O}_{j\cdot}^{\mathrm{T}}\mathcal{O}_{k\cdot}.$$

$\mathcal{O}_{j\cdot}$ 的正交性意味着当 $j\neq k$ 时 $\mathcal{O}_{j\cdot}^{\mathrm{T}}\mathcal{O}_{k\cdot}=0$. 因此，上面的二重和化简得到所需要的结果：

$$\sum_{j=0}^{N-1}\beta_j^2=\sum_{j=0}^{N'-1}(O_j-\alpha_j)^2+\sum_{j=N'}^{N-1}O_j^2.$$

练习[46a]答案 定义 $\boldsymbol{O}$ 的第 j 个元素为 $O_j\equiv\langle X,\mathcal{O}_{j\cdot}\rangle=\mathcal{O}_{j\cdot}^{H}X$. 因此有 $\boldsymbol{O}=\mathcal{O}X$，注意

$$\mathcal{O}X=\begin{bmatrix}\mathcal{O}_{0\cdot}^H\\ \mathcal{O}_{1\cdot}^H\\ \vdots\\ \mathcal{O}_{N-1\cdot}^H\end{bmatrix}X=\begin{bmatrix}\mathcal{O}_{0\cdot}^H X\\ \mathcal{O}_{1\cdot}^H X\\ \vdots\\ \mathcal{O}_{N-1\cdot}^H X\end{bmatrix}=\begin{bmatrix}\langle X,\mathcal{O}_{0\cdot}\rangle\\ \langle X,\mathcal{O}_{1\cdot}\rangle\\ \vdots\\ \langle X,\mathcal{O}_{N-1\cdot}\rangle\end{bmatrix}=\begin{bmatrix}O_0\\ O_1\\ \vdots\\ O_{N-1}\end{bmatrix},$$

这建立了分析方程. 因为由定义 $\mathcal{O}^H\mathcal{O}=I_N$，在分析方程的两边同乘以 $\mathcal{O}^H$ 就可以容易得到综合方程.

练习[46b]答案 首先，注意到，因为 $e^0=1$，

$$F_0=\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t,$$

因为 X_t 为实数，所以上式为实数. 其次，因为 $e^{-i2\pi t}=1$，对所有整数 t，得到

$$F_{N-k}=\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t e^{-i2\pi t(N-k)/N}$$

$$=\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t e^{-i2\pi t}e^{i2\pi tk/N}$$

$$=\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t e^{i2\pi tk/N}$$

$$=\left(\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t e^{-i2\pi tk/N}\right)^*=F_k^*.$$

最后，注意到因为 $e^{-i\pi t}=(e^{-i\pi})^t=(-1)^t$，则有

$$F_{\frac{N}{2}}=\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t e^{-i2\pi t\left(\frac{N}{2}\right)/N}=\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t(-1)^t,$$

上式必须为实值.

练习[47]答案 $\mathcal{F}^H$ 的第 j 行元素为 $\exp(i2\pi tj/N)/\sqrt{N}$，$t=0,\cdots,N-1$。而 $\mathcal{F}$ 的第 k 列元素为 $\exp(-i2\pi tk/N)/\sqrt{N}$，$t=0,\cdots,N-1$，因此 $\mathcal{F}^H\mathcal{F}$ 的第(j,k)个元素为

$$\begin{aligned}A_{j,k} &= \frac{1}{N}\sum_{t=0}^{N-1}e^{i2\pi tj/N}e^{-i2\pi tk/N}\\ &= \frac{1}{N}\sum_{t=0}^{N-1}e^{i2\pi(j-k)t/N}\\ &= \frac{1}{N}\sum_{t=0}^{N-1}z^t,\end{aligned}$$

而

$$z \equiv e^{i2\pi(j-k)/N}.$$

复指数 $\exp(i2\pi x)=\cos(2\pi x)+i\sin(2\pi x)$ 等于 1 当且仅当 x 为整数；因为 $-(N-1)\leqslant j-k\leqslant N-1$，所以仅当 $j=k$ 时 $z=1$，此时上面的最后的和为 N。因此当 $j=k$ 时有 $A_{k,k}=1$，而 $j\neq k$ 时我们有(利用练习[29a]的提示)

$$A_{j,k}=\frac{1}{N}\left(\frac{1-z^N}{1-z}\right)=\frac{1}{N}\left(\frac{1-e^{i2\pi(j-k)}}{1-e^{i2\pi(j-k)/N}}\right)=0,$$

因为 $1-e^{i2\pi(j-k)}=1-\cos[2\pi(j-k)]-i\sin[2\pi(j-k)]=0$。

练习[50a]答案 我们需要证明对于 $0<k<\frac{N}{2}$ 有

$$F_k\mathcal{F}_{k\cdot}+F_{N-k}\mathcal{F}_{N-k\cdot}=2\mathcal{R}(F_k\mathcal{F}_{k\cdot}).$$

首先，注意

$$\begin{aligned}F_{N-k} &\equiv \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t e^{-i2\pi t(N-k)/N}\\ &= \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t e^{i2\pi t}e^{i2\pi tk/N}\\ &= \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1}X_t e^{i2\pi tk/N}\\ &= F_k^*,\end{aligned}$$

其中用到了事实 $\exp(i2\pi t)=1$，因为 t 为整数。其次，注意到，$\mathcal{F}$ 的第$(N-k)$行是 $\mathcal{F}^H_{N-k\cdot}$，因此 $\mathcal{F}_{N-k\cdot}$ 的第 t 个元素为

$$\left(\frac{1}{\sqrt{N}}e^{-i2\pi t(N-k)/N}\right)^*=\frac{1}{\sqrt{N}}e^{-i2\pi tk/N},$$

上式等于 $\mathcal{F}_{k\cdot}^*$ 的第 t 个元素；即 $\mathcal{F}_{N-k\cdot}=\mathcal{F}_{k\cdot}^*$。因此

$$F_k\mathcal{F}_{k\cdot}+F_{N-k}\mathcal{F}_{N-k\cdot}=F_k\mathcal{F}_{k\cdot}+F_k^*\mathcal{F}_{k\cdot}^*=F_k\mathcal{F}_{k\cdot}+(F_k\mathcal{F}_{k\cdot})^*.$$

现在，如果 $z=x+iy$ 是任一复值变量，则 $z+z^*=x+iy+x-iy=2x=2\mathcal{R}(z)$，我们就可以得到结论 $F_k\mathcal{F}_{k\cdot}+F_{N-k}\mathcal{F}_{N-k\cdot}=2\mathcal{R}(F_k\mathcal{F}_{k\cdot})$。

练习[50b]答案 由练习[50a]的结论可知 $D_{\mathcal{F},k}=2\mathcal{R}(F_k\mathcal{F}_{k\cdot})$。$\mathcal{F}$ 的第 k 行是 $\mathcal{F}^H_{k\cdot}$，因此它的第 t 个分量由

$$\frac{e^{i2\pi tk/N}}{\sqrt{N}}=\frac{e^{i2\pi f_k t}}{\sqrt{N}}$$

给出，并且 $F_k\ \mathcal{F}_{k\cdot}$ 的第 t 个分量由

$$\frac{1}{\sqrt{N}}F_k e^{i2\pi f_k t} = \frac{1}{\sqrt{N}}(A_k - iB_k)(\cos(2\pi f_k t) + i\sin(2\pi f_k t))$$

$$= \frac{1}{\sqrt{N}}[A_k\cos(2\pi f_k t) + B_k\sin(2\pi f_k t) + i(A_k\sin(2\pi f_k t) - B_k\cos(2\pi f_k t))]$$

给出，从上面可以得到 $2\,\mathcal{R}(F_k\ \mathcal{F}_{k\cdot})$ 的第 t 个分量，即 $\mathcal{D}_{\mathcal{F},k,t}$ 由

$$\frac{2}{\sqrt{N}}[A_k\cos(2\pi f_k t) + B_k\sin(2\pi f_k t)]$$

给出.

练习[52]答案 注意，对于任何整数 m，$\mathcal{T}^m\boldsymbol{X}$ 的元素可表示为

$$X_{-m \bmod N}, X_{1-m \bmod N}, X_{2-m \bmod N}, \cdots, X_{N-1-m \bmod N},$$

因此 $\mathcal{T}^m\boldsymbol{X}$ 的第 k 个傅里叶系数由

$$\begin{aligned}
\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1} X_{t-m \bmod N}e^{-i2\pi tk/N} &= \frac{1}{\sqrt{N}}\sum_{t=-m}^{N-1-m} X_{t \bmod N}e^{-i2\pi(t+m)k/N}\\
&= \frac{e^{-i2\pi mk/N}}{\sqrt{N}}\sum_{t=-m}^{N-1-m} X_{t \bmod N}e^{-i2\pi tk/N}\\
&= \frac{e^{-i2\pi mk/N}}{\sqrt{N}}\sum_{t=0}^{N-1} X_{t \bmod N}e^{-i2\pi tk/N}\\
&= \frac{e^{-i2\pi mk/N}}{\sqrt{N}}\sum_{t=0}^{N-1} X_t e^{-i2\pi tk/N}\\
&= e^{-i2\pi mk/N}F_k
\end{aligned}$$

给出，其中我们用到了下面两个事实：(i)无穷序列 $\{X_{t \bmod N}e^{-i2\pi tk/N}:\ t=\cdots,\ -1,\ 0,\ 1,\ \cdots\}$ 是以 N 为周期的周期序列，因此任意 N 个邻接元素的和都等于任何另外的 N 个邻接元素的和；(ii) $X_{t \bmod N} = X_t$，$t=0,\ \cdots,\ N-1$.

练习[58]答案 令 $r_1=0$，并且定义 $r_j = r_{j-1} + N/2^{j-1}$，对 $j=2,\ 3,\ \cdots,\ J$，其中 r_j 用来标示 $\mathcal{W}$ 的特殊行(例如：当 $N=16$，$J=4$ 时，r_j 标示行为 $r_1=0$，$r_2=8$，$r_3=12$，$r_4=14$)．定义 $\mathcal{W}$ 的行 $r_j(j=1,\ 2,\ \cdots,\ J)$ 为

$$\mathcal{W}_{r_j\cdot}^{T} = \Big[\underbrace{-\frac{1}{2^{j/2}},\cdots,-\frac{1}{2^{j/2}}}_{2^{j-1}\text{个}},\underbrace{\frac{1}{2^{j/2}},\cdots,\frac{1}{2^{j/2}}}_{2^{j-1}\text{个}},\underbrace{0,\cdots,0}_{N-2^j\text{个}0}\Big],$$

因为

$$\|\mathcal{W}_{r_j\cdot}\|^2 = 2^{j-1}\left(-\frac{1}{2^{j/2}}\right)^2 + 2^{j-1}\left(\frac{1}{2^{j/2}}\right)^2 = \frac{2^{j-1}}{2^j} + \frac{2^{j-1}}{2^j} = 1,$$

所以，所有 $\mathcal{W}_{r_j\cdot}$ 范数均为 1；另外，如果 $j<j'$，我们有

$$\begin{aligned}
\langle \mathcal{W}_{r_j\cdot}, \mathcal{W}_{r_{j'}\cdot}\rangle &= 2^{j-1}\left(-\frac{1}{2^{j/2}}\right)\left(-\frac{1}{2^{j'/2}}\right) + 2^{j-1}\left(\frac{1}{2^{j/2}}\right)\left(-\frac{1}{2^{j'/2}}\right)\\
&= \frac{2^{j-1}}{2^{(j+j')/2}} - \frac{2^{j-1}}{2^{(j+j')/2}}\\
&= 0,
\end{aligned}$$

因此$\mathcal{W}_{r_j \cdot}$是规范正交的．接着定义$\mathcal{W}$的剩余行如下：

$$\mathcal{W}_{r_j+k \cdot} = \mathcal{T}^{2^j k}\ \mathcal{W}_{r_j \cdot},\quad j=1,\cdots,J-1, k=1,\cdots,\frac{N}{2^j}-1;$$

$$\mathcal{W}_{N-1 \cdot} = \left[\frac{1}{2^{J/2}},\cdots,\frac{1}{2^{J/2}}\right]^{\mathrm{T}},$$

因为

$$\|\ \mathcal{W}_{r_j+k \cdot}\ \|^2 = \|\ \mathcal{T}^{2^j k}\ \mathcal{W}_{r_j \cdot}\ \|^2 = \|\ \mathcal{W}_{r_j \cdot}\ \|^2 = 1$$

和

$$\|\ \mathcal{W}_{N-1 \cdot}\ \|^2 = N\left(\frac{1}{2^{J/2}}\right)^2 = \frac{2^J}{2^J} = 1,$$

所以$\mathcal{W}$的每行具有 1 范数．因为

$$\langle \mathcal{W}_{N-1 \cdot}, \mathcal{W}_{r_j+k \cdot}\rangle = \frac{1}{2^{J/2}}\left[2^{j-1}\left(-\frac{1}{2^{j/2}}\right)+2^{j-1}\left(\frac{1}{2^{j/2}}\right)\right]=0$$

对于$j=1$，…，J和$k=0$，…，$\frac{N}{2^j}-1$成立．由此得出$\mathcal{W}_{N-1 \cdot}$与$\mathcal{W}$的剩余$N-1$行是规范正交的．最后，我们必须证明任选其他的两行都是正交的，即

$$\langle \mathcal{W}_{r_j+k \cdot}, \mathcal{W}_{r_{j'}+k' \cdot}\rangle = 0.$$

现在

$$\begin{aligned}\langle \mathcal{W}_{r_j+k \cdot}, \mathcal{W}_{r_{j'}+k' \cdot}\rangle &= \langle \mathcal{T}^{2^j k}\ \mathcal{W}_{r_j \cdot}, \mathcal{T}^{2^{j'} k'}\ \mathcal{W}_{r_{j'} \cdot}\rangle \\ &= \mathcal{W}_{r_j \cdot}^{\mathrm{T}}\ \mathcal{T}^{-2^j k}\ \mathcal{T}^{2^{j'} k'}\ \mathcal{W}_{r_{j'} \cdot} \\ &= \mathcal{W}_{r_j \cdot}^{\mathrm{T}}\ \mathcal{T}^{2^{j'} k'-2^j k}\ \mathcal{W}_{r_{j'} \cdot}.\end{aligned}$$

我们能够选择$j<j'$使$j'=j+\delta$，$\delta>0$．然后$2^{j'}k'-2^j k=2^j(2^\delta k'-k)$，此式说明平移是$2^j$的整数倍；然而，加到$\mathcal{W}_{r_{j'} \cdot}$上的量所作的前后平移将导致$\mathcal{W}_{r_j \cdot}^{\mathrm{T}}$的一串等值(公共值为 0，或$1/2^{j'/2}$，或者$-1/2^{j'/2}$)．因此内积总得到为 0，得证．

练习[66]答案 回忆$\mathcal{V}_J$是正交矩阵$\mathcal{W}$的最后行，$\mathcal{W}_j$是$\mathcal{W}$的另外$N/2^j$行组成的$N/2^j \times N$矩阵．由正交性得$\mathcal{W}_j\ \mathcal{V}_J^{\mathrm{T}}=\mathbf{0}$，其中$\mathbf{0}$是$N/2^j$维零向量．因此，对于$1 \leqslant j \leqslant J$，

$$\mathcal{D}_j^{\mathrm{T}}\ \mathcal{V}_J^{\mathrm{T}} = \mathbf{W}_j^{\mathrm{T}}\ \mathcal{W}_j\ \mathcal{V}_J^{\mathrm{T}} = 0.$$

练习[97]说明$\mathcal{V}_J$的所有元素为$1/\sqrt{N}$，因此上式左边与$\mathcal{D}_j$的元素之和成比例，因此我们得到$\mathcal{D}_j$的元素的样本均值为 0．

练习[69]答案 假设$\{h_l'$：$l=0$，…，$L-1\}$是具有奇数宽度L的小波滤波器．规范正交性条件(式(69c))要求，对于所有的非 0 整数我们必须有

$$\sum_{l=0}^{L-1} h_l' h_{l+2n}' = 0$$

(这里$h_l' \equiv 0$，对$l<0$和$l \geqslant L$)．如果L为奇数，则$L-1$为偶数，因此令$n=(L-1)/2$得到(因为$h_0' \neq 0$和$h_{L-1}' \neq 0$)

$$\sum_{l=0}^{L-1} h_l' h_{l+L-1}' = h_0' h_{L-1}' \neq 0,$$

这违背了式(69c)．

练习[70]答案 因为 $\{h\star h_j\}$ 是 $\{h_j\}$ 的自相关序列，所以 $\{h\star h_j\}$ 的离散傅里叶变换由 $\{h_j\}$ 的离散傅里叶的模平方给出，即 $\mathcal{H}(f)=|H(f)|^2$（见式(36e)）．使用 $\exp(-\mathrm{i}\pi)=-1$ 和对非零偶数 $j h\star h_j=0$ 的假设，我们有

$$\mathcal{H}(f)=\sum_{j=-\infty}^{\infty}h\star h_j\mathrm{e}^{-\mathrm{i}2\pi fj}=1+\sum_{j=\pm1,\pm3,\cdots}h\star h_j\mathrm{e}^{-\mathrm{i}2\pi fj},$$

$$\begin{aligned}\mathcal{H}\left(f+\frac{1}{2}\right)&=\sum_{j=-\infty}^{\infty}h\star h_j\mathrm{e}^{-\mathrm{i}2\pi\left(f+\frac{1}{2}\right)j}\\&=1+\sum_{j=\pm1,\pm3,\cdots}[\mathrm{e}^{-\mathrm{i}\pi}]^j h\star h_j\mathrm{e}^{-\mathrm{i}2\pi fj}\\&=1+\sum_{j=\pm1,\pm3,\cdots}[-1]^j h\star h_j\mathrm{e}^{-\mathrm{i}2\pi fj}\\&=1-\sum_{j=\pm1,\pm3,\cdots}h\star h_j\mathrm{e}^{-\mathrm{i}2\pi fj},\end{aligned}$$

因为对所有的奇数 j，有 $[-1]^j=-1$．产生 $\mathcal{H}(f)+\mathcal{H}\left(f+\frac{1}{2}\right)=2$.

换句话说，因为 $\{h\star h_j\}\leftrightarrow\mathcal{H}(\cdot)$，所以用练习[23b]得到

$$h\star h_{2n}=\frac{1}{2}\int_{-1/2}^{1/2}\left[\mathcal{H}\left(\frac{f}{2}\right)+\mathcal{H}\left(\frac{f}{2}+\frac{1}{2}\right)\right]\mathrm{e}^{\mathrm{i}2\pi fn}\,\mathrm{d}f.$$

上面说明 $\mathcal{H}\left(\frac{f}{2}\right)+\mathcal{H}\left(\frac{f}{2}+\frac{1}{2}\right)$ 定义了一个函数，该函数的逆离散傅里叶变换是指标 $n=0$ 时为 2 其他时为 0 的序列；然而，这个序列的离散傅里叶变换是处处恒等于 2 的函数，即 $\mathcal{H}\left(\frac{f}{2}\right)+\mathcal{H}\left(\frac{f}{2}+\frac{1}{2}\right)=2$.

练习[72]答案 因为对于所有整数 t，$\mathrm{e}^{\mathrm{i}2\pi t}=1$ 和对于所有的 f，有 $\mathcal{H}(f)+\mathcal{H}\left(f+\frac{1}{2}\right)=2$，我们有

$$\begin{aligned}h^\circ\star h^\circ_{2t}&=\frac{1}{N}\left(\sum_{k=0}^{\frac{N}{2}-1}\mathcal{H}\left(\frac{k}{N}\right)\mathrm{e}^{\mathrm{i}2\pi(2t)k/N}+\sum_{k=0}^{\frac{N}{2}-1}\mathcal{H}\left(\frac{k}{N}+\frac{1}{2}\right)\mathrm{e}^{\mathrm{i}2\pi(2t)\left(\frac{k}{N}+\frac{1}{2}\right)}\right)\\&=\frac{1}{N}\sum_{k=0}^{\frac{N}{2}-1}\left(\mathcal{H}\left(\frac{k}{N}\right)+\mathcal{H}\left(\frac{k}{N}+\frac{1}{2}\right)\right)\mathrm{e}^{\mathrm{i}2\pi(2t)k/N}\\&=\frac{2}{N}\sum_{k=0}^{\frac{N}{2}-1}\mathrm{e}^{\mathrm{i}2\pi(2t)k/N}.\end{aligned}$$

为了得到想要的结果，注意

$$\sum_{k=0}^{\frac{N}{2}-1}\mathrm{e}^{\mathrm{i}2\pi(2t)k/N}=\begin{cases}N/2, & t=0;\\0, & t=1,\cdots,\frac{N}{2}-1,\end{cases}$$

当 $t=0$ 上式是明显的，而对于其他的 t，应用

$$\sum_{k=0}^{\frac{N}{2}-1}z^k=\frac{1-z^{N/2}}{1-z},$$

而 $z=\mathrm{e}^{\mathrm{i}2\pi(2t)/N}$，所以

$$z^{N/2}=\mathrm{e}^{\mathrm{i}2\pi t}=1.$$

练习[76a]答案 因为 $g_l=(-1)^{l+1}h_{L-l-1}$，我们有

$$
\begin{aligned}
G(f) &\equiv \sum_{l=0}^{L-1} g_l e^{-i2\pi fl} \\
&= \sum_{l=0}^{L-1}(-1)^{l+1}h_{L-l-1}e^{-i2\pi fl} \\
&= \sum_{l=0}^{L-1}(-1)^{L-l}h_l e^{-i2\pi f(L-1-l)} \\
&= \sum_{l=0}^{L-1}(-1)^{l}h_l e^{-i2\pi f(L-1-l)},
\end{aligned}
$$

其中我们用到了事实：$(-1)^{L-l}=(-1)^L(-1)^{-l}=(-1)^l$，因为 L 是偶数(因此 $(-1)^L=1$)和 $(-1)^{-l}=(-1)^l$. 注意到 $e^{-i\pi}=-1$，因此 $(-1)^l=e^{-i\pi l}$，我们有

$$
\begin{aligned}
G(f) &= \sum_{l=0}^{L-1} e^{-i\pi l}h_l e^{-i2\pi f(L-1-l)} \\
&= e^{-i2\pi f(L-1)}\sum_{l=0}^{L-1}h_l e^{-i2\pi\left(\frac{1}{2}-f\right)l} \\
&= e^{-i2\pi f(L-1)}H\left(\frac{1}{2}-f\right).
\end{aligned}
$$

因为

$$
\mathcal{G}(f)\equiv|G(f)|^2=\left|e^{-i2\pi f(L-1)}H\left(\frac{1}{2}-f\right)\right|^2=\left|H\left(\frac{1}{2}-f\right)\right|^2=\mathcal{H}\left(\frac{1}{2}-f\right),
$$

第二部分容易得到.

练习[76b]答案 由练习[76a]我们有

$$
\sum_{l=0}^{L-1} g_l = G(0) = H\left(\frac{1}{2}\right).
$$

在式(69d)中，令 $f=0$ 得到 $|H(0)|^2+\left|H\left(\frac{1}{2}\right)\right|^2=2$；然而，由式(69a)有

$$
H(0)=\sum_{l=0}^{L-1}h_l=0,
$$

所以 $\left|H\left(\frac{1}{2}\right)\right|^2=\mathcal{H}\left(\frac{1}{2}\right)=2$，从而 $H\left(\frac{1}{2}\right)=\pm\sqrt{2}$，建立了练习的第一部分.

因为对 $\{g_l\}$ 的平方增益函数 $\mathcal{G}(\cdot)$ 满足 $\mathcal{G}(f)+\mathcal{G}\left(f+\frac{1}{2}\right)=2$（这是式(76b)的第一部分），我们从下述的公式(69d)知道 $\{g_l\}$ 必须满足规范正交性质，从而得到练习的第二部分. 另一方面，我们可以如下直接建立规范正交性质的第二部分. 使用 $g_l\equiv(-1)^{l+1}h_{L-l-1}$，我们有

$$
\sum_{l=0}^{L-1}g_l^2=\sum_{l=0}^{L-1}(-1)^{2l+2}h_{L-l-1}^2=\sum_{l=0}^{L-1}h_l^2=1,
$$

因为小波滤波器有单位能量(式(69a)). 第二，对于所有非 0 整数均有

$$
\begin{aligned}
\sum_{l=0}^{L-1}g_l g_{l+2n} &= \sum_{l=0}^{L-1}(-1)^{l+1}h_{L-l-1}(-1)^{l+2n+1}h_{L-1-(l+2n)} \\
&= \sum_{l=0}^{L-1}(-1)^{2(l+n+1)}h_{L-1-l}h_{L-1-(l+2n)} \\
&= \sum_{l=0}^{L-1}h_l h_{l-2n}=0,
\end{aligned}
$$

因为小波滤波器与它的偶平移是正交的(式(69c)).

练习[77]答案 对于所有整数 n 我们有

$$\sum_{l=-\infty}^{\infty} g_l h_{l+2n} = \sum_{l=-\infty}^{\infty} (-1)^{l+1} h_{L-1-l} h_{l+2n} \equiv S,$$

如果由 l 用变量代替 $m \equiv L-1-l-2n$(所以 $l=L-1-m-2n$)，我们有

$$\begin{aligned}
S &= \sum_{m=-\infty}^{\infty} (-1)^{(L-1-m-2n)+1} h_{L-1-(L-1-m-2n)} h_{(L-1-m-2n)+2n} \\
&= \sum_{m=-\infty}^{\infty} (-1)^{L-m-2n} h_{m+2n} h_{L-1-m} \\
&= \sum_{m=-\infty}^{\infty} (-1)^{-m} h_{m+2n} h_{L-1-m} \quad (\text{因为 } L-2n \text{ 为偶数}) \\
&= \sum_{m=-\infty}^{\infty} (-1)^{m} h_{m+2n} h_{L-1-m} \\
&= -\sum_{m=-\infty}^{\infty} (-1)^{m+1} h_{L-1-m} h_{m+2n} = -S,
\end{aligned}$$

由此得到所需的结论 $S=0$.

另一方面这里是和练习[78]解法紧密相关的频域方法．因此 g_l 为实值．从式(36d)可知互相关

$$g \star h_n \equiv \sum_{l=-\infty}^{\infty} g_l h_{l+n}$$

的离散傅里叶变换由 $G^*(f)H(f)$ 给出．令 $b_n \equiv g \star h_{2n}$．我们需要证明对所有的 n 有 $b_n=0$．由练习[23b]得出 $\{b_l\}$ 的离散傅里叶变换由

$$B(f) = \frac{1}{2}\left[G^*\left(\frac{f}{2}\right)H\left(\frac{f}{2}\right) + G^*\left(\frac{f}{2}+\frac{1}{2}\right)H\left(\frac{f}{2}+\frac{1}{2}\right)\right]$$

得出．式(76a)说明 $G(f)=e^{-i2\pi f(L-1)}H\left(\frac{1}{2}-f\right)$，因此

$$G^*(f) = e^{i2\pi f(L-1)}H^*\left(\frac{1}{2}-f\right) = e^{i2\pi f(L-1)}H\left(f-\frac{1}{2}\right) = e^{i2\pi f(L-1)}H\left(f+\frac{1}{2}\right),$$

因为对于实值序列，$H^*(f)=H(-f)$(见练习[22b]，并且 $H(\cdot)$ 是单位周期函数(参见练习[22a])．因为 $G(\cdot)$ 也具有单位周期性，所以上面方括号里面的项变为

$$\begin{aligned}
&e^{i\pi f(L-1)}H\left(\frac{f}{2}+\frac{1}{2}\right)H\left(\frac{f}{2}\right) + e^{i\pi(f+1)(L-1)}H\left(\frac{f}{2}\right)H\left(\frac{f}{2}+\frac{1}{2}\right) \\
&= e^{i\pi f(L-1)}\left[H\left(\frac{f}{2}+\frac{1}{2}\right)H\left(\frac{f}{2}\right) + e^{i\pi(L-1)}H\left(\frac{f}{2}\right)H\left(\frac{f}{2}+\frac{1}{2}\right)\right] \\
&= e^{i\pi f(L-1)}\left[H\left(\frac{f}{2}+\frac{1}{2}\right)H\left(\frac{f}{2}\right) - H\left(\frac{f}{2}\right)H\left(\frac{f}{2}+\frac{1}{2}\right)\right] = 0,
\end{aligned}$$

因为 $e^{i\pi(L-1)}=[e^{i\pi}]^{L-1}=[-1]^{L-1}=-1$(回忆 $L-1$ 必定是奇数)．因此对于所有 f，$B(f)=0$，这告诉我们对于所有的 n，$b_n=0$.

练习[78]答案 由

$$g^{\circ} \star h^{\circ}_{2n} = \frac{1}{N}\sum_{k=0}^{N-1} G^{*}\left(\frac{k}{N}\right)H\left(\frac{k}{N}\right)e^{i4\pi nk/N}$$

$$= \frac{1}{N}\left(\sum_{k=0}^{\frac{N}{2}-1} G^{*}\left(\frac{k}{N}\right)H\left(\frac{k}{N}\right)e^{i4\pi nk/N} + \sum_{k=0}^{\frac{N}{2}-1} G^{*}\left(\frac{k}{N}+\frac{1}{2}\right)H\left(\frac{k}{N}+\frac{1}{2}\right)e^{i4\pi n\left(\frac{k}{N}+\frac{1}{2}\right)}\right)$$

$$= \frac{1}{N}\sum_{k=0}^{\frac{N}{2}-1}\left[G^{*}\left(\frac{k}{N}\right)H\left(\frac{k}{N}\right)+G^{*}\left(\frac{k}{N}+\frac{1}{2}\right)H\left(\frac{k}{N}+\frac{1}{2}\right)\right]e^{i4\pi nk/N}$$

得到第一项展开式. 为了得到第二项展开式，首先注意下述的事实：

$$G(f) = e^{-i2\pi f(L-1)}H\left(\frac{1}{2}-f\right) \qquad \text{由式(76a)}$$

$$G^{*}(f) = e^{i2\pi f(L-1)}H^{*}\left(\frac{1}{2}-f\right) = e^{i2\pi f(L-1)}H\left(f-\frac{1}{2}\right)$$

$$G^{*}\left(f+\frac{1}{2}\right) = e^{i2\pi(f+1/2)(L-1)}H(f) = -e^{i2\pi f(L-1)}H(f),$$

其中我们使用了事实 $e^{i\pi(L-1)} = (e^{-i\pi})^{L-1} = (-1)^{L-1} = -1$，因此 L 为偶数. 因此

$$G^{*}\left(\frac{k}{N}\right)H\left(\frac{k}{N}\right)+G^{*}\left(\frac{k}{N}+\frac{1}{2}\right)H\left(\frac{k}{N}+\frac{1}{2}\right)$$

$$= e^{i2\pi\frac{k}{N}(L-1)}H\left(\frac{k}{N}-\frac{1}{2}\right)H\left(\frac{k}{N}\right)-e^{i2\pi\frac{k}{N}(L-1)}H\left(\frac{k}{N}\right)H\left(\frac{k}{N}+\frac{1}{2}\right)=0,$$

因为 $H(\cdot)$是单位周期函数，所以 $H\left(\frac{k}{N}-\frac{1}{2}\right)=H\left(\frac{k}{N}+\frac{1}{2}\right)$.

练习[91]答案 令$\{a_l\}\leftrightarrow A(\cdot)$表示新的滤波器. 因为

$$a_l = \begin{cases} h_{l/(m+1)}, & l \text{ 是 } m+1 \text{ 的整数倍；} \\ 0, & \text{其他，} \end{cases}$$

如所需要的，我们有

$$A(f) = \sum_{l=0}^{(m+1)(L-1)} a_l e^{-i2\pi fl}$$

$$= \sum_{l=0}^{L-1} a_{(m+1)l} e^{-i2\pi f(m+1)l}$$

$$= \sum_{l=0}^{L-1} h_l e^{-i2\pi([m+1]f)l}$$

$$= H([m+1]f).$$

练习[92]答案 因为$\{h_{2,l}\}\leftrightarrow H_2(\cdot)$和 $H_2(f)=H(2f)G(f)$，我们有 $H_2(0)=H(0)G(0)$. 因为 $H(0)=\sum_l h_l=0$,我们得到所求

$$H_2(0) = \sum_{l=0}^{L_2-1} h_{2,l} = 0.$$

当 $N \geq L_2$ 时，$\{h_{2,l}\}$的元素形成离散傅里叶变换矩阵$\mathcal{W}$的其中一行的元素.因为矩阵$\mathcal{W}$是规范正交的，每行必须有单位能量，因此$\{h_{2,l}\}$也如此．另一方面，用帕塞瓦尔(式(35c))定理得到

$$\begin{aligned}\sum_{l=-\infty}^{\infty} h_{2,l}^2 &= \int_{-1/2}^{1/2} \mathcal{H}_2(f)\mathrm{d}f \\ &= \int_{-1/2}^{1/2} \mathcal{H}(2f)\,\mathcal{G}(f)\mathrm{d}f \\ &= \frac{1}{2}\int_{-1}^{1} \mathcal{H}(f')\,\mathcal{G}\left(\frac{f'}{2}\right)\mathrm{d}f'.\end{aligned}$$

就像证明练习[23b]一样，我们把最后一个积分分成三部分并且用积分变量代换得到

$$\begin{aligned}\sum_{l=-\infty}^{\infty} h_{2,l}^2 &= \frac{1}{2}\int_{-1/2}^{1/2} \mathcal{H}(f)\,\mathcal{G}\left(\frac{f}{2}\right)+\mathcal{H}(f)\,\mathcal{G}\left(\frac{f}{2}+\frac{1}{2}\right)\mathrm{d}f \\ &= \int_{-1/2}^{1/2} \mathcal{H}(f)\mathrm{d}f = \sum_{l=-\infty}^{\infty} h_l^2 = 1,\end{aligned}$$

其中我们用到了$\mathcal{G}\left(\frac{f}{2}\right)+\mathcal{G}\left(\frac{f}{2}+\frac{1}{2}\right)=2$ 和再次用到帕塞瓦尔定理.

为了说明$\{h_{2,l}\}$对于偶平移不能是正交的，假设结论成立．因为它具有单位能量，所以它的平方增益函数必须满足正交性条件$\mathcal{H}_2\left(\frac{f}{2}\right)+\mathcal{H}_2\left(\frac{f}{2}+\frac{1}{2}\right)=2$，对于所有的 f 都成立．然而，对于 $f=0$ 和使用$\mathcal{H}_2(f)=\mathcal{H}(2f)\mathcal{G}(f)$，我们有

$$\mathcal{H}_2(0)+\mathcal{H}_2\left(\frac{1}{2}\right)=\mathcal{H}(0)\,\mathcal{G}(0)+\mathcal{H}(1)\,\mathcal{G}\left(\frac{1}{2}\right)=0,$$

因为$\mathcal{H}(1)=\mathcal{H}(0)=0$，导致矛盾．因此$\{h_{2,l}\}$对于偶平移不能是正交的.

第二层哈尔小波滤波器由

$$\{h_{2,l}\} \equiv \{g * h_l^{\uparrow}\} = \left\{\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right\} * \left\{\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}}\right\} = \left\{\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right\}$$

给出．这个滤波器和为 0 并且具有单位能量，但是

$$\sum_{l=0}^{1} h_{2,l}h_{2,l+2} = \frac{1}{2}\cdot\left(-\frac{1}{2}\right)+\frac{1}{2}\cdot\left(-\frac{1}{2}\right)=-\frac{1}{2}\neq 0,$$

因此式(69c)不满足.

练习[95]答案 注意$\mathcal{D}_k=\mathcal{W}_k^T\mathbf{W}_k$．所以$\|\mathcal{D}_k\|^2=\mathbf{W}_k^T\mathcal{W}_k\mathcal{W}_k^T\mathbf{W}_k$．因为$\mathcal{W}_k=\mathcal{B}_k\mathcal{A}_{k-1}\cdots\mathcal{A}_1$，我们有

$$\begin{aligned}\mathcal{W}_k\mathcal{W}_k^T &= \mathcal{B}_k\mathcal{A}_{k-1}\cdots\mathcal{A}_1\mathcal{A}_1^T\cdots\mathcal{A}_{k-1}^T\mathcal{B}_k^T \\ &= \mathcal{B}_k\mathcal{A}_{k-1}\cdots\mathcal{A}_2 I_{N_1}\mathcal{A}_2^T\cdots\mathcal{A}_{k-1}^T\mathcal{B}_k^T \\ &= \mathcal{B}_k\mathcal{A}_{k-1}\cdots\mathcal{A}_2\mathcal{A}_2^T\cdots\mathcal{A}_{k-1}^T\mathcal{B}_k^T \\ &\;\;\vdots \\ &= \mathcal{B}_k\mathcal{B}_k^T \\ &= I_{N_k}.\end{aligned}$$

因此

$$\|\mathcal{D}_k\|^2=\mathbf{W}_k^T I_{N_k}\mathbf{W}_k=\mathbf{W}_k^T\mathbf{W}_k=\|\mathbf{W}_k\|^2.$$

类似地，$\|\mathcal{S}_j\|^2=\|\mathbf{V}_j\|^2$，并且结论由式(95d)得到.

练习[96]答案 练习[28a]的结果说明了等价滤波器的长度是 j 个独立的滤波器的长度之和再减去 $j-1$ 得到．现在，第 k 个滤波器($k=1,\cdots,j$)有 2^{k-1} 个0，紧接着 $L-1$ 个非0元素，因此有 $(L-1)(2^{k-1}-1)$ 个0元素和 L 个非0元素，总共有 $(L-1)2^{k-1}+1$ 个元素．因此，因为

$$\sum_{k=1}^{j} 2^{k-1} = 2^j - 1,$$

滤波器大小为

$$\sum_{k=1}^{j}\left[(L-1)2^{k-1}+1\right]-(j-1) = (2^j-1)(L-1)+1 \equiv L_j.$$

级联滤波器的传递函数是单独的传递函数之积．第一个滤波器具有传递函数为 $G(\cdot)$．第二个滤波器由对第一个进行上抽样得到，因此练习[91]说明了它的传递函数为 $G(2f)$；并且第三个滤波器由对第二个进行上抽样得到，因此它的传递函数为 $G(4f)$；重复这个过程，首先的 $j-1$ 个滤波器的第 k 个的传递函数为 $G(2^{k-1}f)$．类似地，对于第 j 个滤波器的传递函数为 $H(2^{j-1}f)$，因此所有这些的乘积就得到了式(96b)给出的 $H_j(f)$．

练习[97]答案 离散傅里叶变换矩阵 $\mathcal{W}$ 的最后一行 $\mathcal{V}_J$ 由

$$\left[g^{\circ}_{J,N-1}, g^{\circ}_{J,N-2}, \cdots, g^{\circ}_{J,1}, g^{\circ}_{J,0}\right]$$

给出．考虑

$$\sum_{l=0}^{N-1}\left(g^{\circ}_{J,l}-\frac{1}{\sqrt{N}}\right)^2 = \sum_{l=0}^{N-1}(g^{\circ}_{J,l})^2 - \frac{2}{\sqrt{N}}\sum_{l=0}^{N-1} g^{\circ}_{J,l}+1.$$

为了使得上式等于0，从而得到对于 $l=0,\cdots,N-1$ 有 $g^{\circ}_{J,l}=\frac{1}{\sqrt{N}}$．为此，注意因为 $\mathcal{V}_J$ 是正交阵的一行，所以必须有

$$\sum_{l=0}^{N-1}(g^{\circ}_{J,l})^2 = 1.$$

因此

$$\sum_{l=0}^{N-1}\left(g^{\circ}_{J,l}-\frac{1}{\sqrt{N}}\right)^2 = 2-\frac{2}{\sqrt{N}}\sum_{l=0}^{N-1} g^{\circ}_{J,l};$$

然而

$$\sum_{l=0}^{N-1} g^{\circ}_{J,l} = \sum_{l=0}^{L_J-1} g_{J,l} = G_J(0),$$

从而

$$\sum_{l=0}^{N-1}\left(g^{\circ}_{J,l}-\frac{1}{\sqrt{N}}\right)^2 = 2-\frac{2G_J(0)}{\sqrt{N}}.$$

回忆

$$G_J(f) = \prod_{l=0}^{J-1} G(2^l f),$$

我们得到

$$G_J(0) = \prod_{l=0}^{J-1} G(0) = \prod_{l=0}^{J-1}\left(\sum_{l=0}^{L-1} g_l\right) = 2^{J/2}.$$

因为 $N=2^J$，我们得到 $G_J(0)=\sqrt{N}$，这就达到了我们的目的．

因为 $\mathcal{V}_J=\left[\frac{1}{\sqrt{N}},\cdots,\frac{1}{\sqrt{N}}\right]$，由此得到

$$W_{N-1} = \mathcal{V}_J \mathbf{X} = \frac{1}{\sqrt{N}}\sum_{t=0}^{N-1} X_t = \overline{X}\sqrt{N}.$$

因为 $\mathbf{V}_J$ 是定义 1×1 矩阵，它的单个元素是 W_{N-1}，那么 $\mathcal{S}_J=\mathcal{V}_J^T\mathbf{V}_J$ 必定为 N 维列向量，其元素均等于 $\overline{X}$．

练习[102a]答案 比较式(95e)和式(96e)可知，我们能够用$\{g_{j-1,l}\}$和滤波器$\{a_l\}$的卷积构造$\{h_{j,l}\}$，该滤波器的脉冲相应序列为

$$h_0, \underbrace{0,\cdots,0}_{2^{j-1}-1\text{个}0}, h_1, \underbrace{0,\cdots,0}_{2^{j-1}-1\text{个}0}, \cdots, h_{L-2}, \underbrace{0,\cdots,0}_{2^{j-1}-1\text{个}0}, h_{L-1};$$

即，我们有

$$a_l \equiv \begin{cases} h_{l/2^{j-1}}, & l = 0, 2^{j-1}, \cdots, 2^{j-1}(L-1); \\ 0, & \text{其他}. \end{cases}$$

因此，$h_{j,l} = \sum_{m=-\infty}^{\infty} g_{j-1,m} a_{l-m}$. 作代换 $l-m=2^{j-1}k$ 就得到第一个结论：

$$h_{j,l} = \sum_{k=-\infty}^{\infty} h_k g_{j-1,l-2^{j-1}k}$$

$$= \sum_{k=0}^{L-1} h_k g_{j-1,l-2^{j-1}k}, \quad l = 0, \cdots, L_j - 1$$

(为了构造$\{g_{j,l}\}$，我们仅需要在定义$\{a_l\}$时用g_l代替h_l).

对于第二个结果，从练习[91]知$\{a_l\} \leftrightarrow A(\cdot)$，其中$A(f) \equiv H(2^{j-1}f)$. 因此，从式(36a)和式(36b)可知，$\{g_{j-1,l}\}$和$\{a_l\}$的卷积$\{h_{j,l}\}$的傅里叶变换由

$$\{h_{j,l}\} \leftrightarrow G_{j-1}(f) H(2^{j-1}f) = H_j(f)$$

具体化. (类似的讨论得到$\{g_{j,l}\} \leftrightarrow G_{j-1}(f)G(2^{j-1}f) = G_j(f)$.)

练习[102b]答案 令$\{a_l\}$在$\{h_{j-1,l}\}$的元素之间插入单个0形成的滤波器：

$$a_l \equiv \begin{cases} h_{j-1,l/2}, & l = 0, 2, \cdots, 2(L_{j-1}-1); \\ 0, & \text{其他}. \end{cases}$$

首先我们证明$\{a_l\}$和$\{g\}$的卷积产生$\{h_{j,l}\} \leftrightarrow H_j(\cdot)$，其中，从式[96]知道

$$H_j(f) = H(2^{j-1}f) \prod_{l=0}^{j-2} G(2^l f).$$

为看到此，注意$\{h_{j-1,l}\} \leftrightarrow H_{j-1}(\cdot)$，其中

$$H_{j-1}(f) = H(2^{j-2}f) \prod_{l=0}^{j-3} G(2^l f).$$

使用练习[91]，对$\{a_l\}$的传递函数定义为

$$H_{j-1}(2f) = H(2^{j-2}[2f]) \prod_{l=0}^{j-3} G(2^l[2f])$$

$$= H(2^{j-1}f) \prod_{l=1}^{j-2} G(2^l f),$$

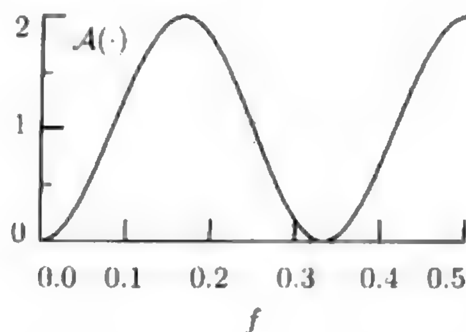


图 519 对滤波器$\{a_0=-1/\sqrt{2},\ a_1=0,\ a_2=0,\ a_3=-1/\sqrt{2}\}$的平方增益函数$\mathcal{A}(\cdot)$

因此$\{a_l\}$和$\{g_l\}$的卷积的传递函数由

$$H_{j-1}(2f)G(f)=\left[H(2^{j-1}f)\prod_{l=1}^{j-2}G(2^lf)\right]G(f)=H(2^{j-1}f)\prod_{l=0}^{j-2}G(2^lf)=H_j(f)$$

给出．因之我们能够写

$$h_{j,l}=\sum_{m=-\infty}^{\infty}g_m a_{l-m}.$$

作变量代换 $l-m=2k$ 得到所求

$$h_{j,l}=\sum_{k=-\infty}^{\infty}g_{l-2k}h_{j-1,k}=\sum_{k=0}^{L_{j-1}-1}g_{l-2k}h_{j-1,k},\quad l=0,\cdots,L_j-1.$$

(我们能够用类似的方法得到$\{g_{j,l}\}$的公式.)

练习[103]答案 因为

$$\sum_{l=-\infty}^{\infty}g_{j,l}=\sum_{l=0}^{2^j-1}\frac{1}{2^{j/2}}=2^{j/2}\quad 而\quad \sum_{l=-\infty}^{\infty}g_{j,l}^2=\sum_{l=0}^{2^j-1}\frac{1}{2^j}=1,$$

利用式(103)就得到所求

$$\text{width}_a\{g_{j,l}\}=\frac{\left(\sum_{l=-\infty}^{\infty}g_{j,l}\right)^2}{\sum_{l=-\infty}^{\infty}g_{j,l}^2}=2^j.$$

练习[105a]答案 因为

$$\sum_{l=0}^{3}a_l=0,\quad \sum_{l=0}^{3}a_l^2=1\quad 和\quad \sum_{l=0}^{3}a_la_{l+2n}=0$$

对于所有的非 0 整数 n 成立，$\{a_l\}$满足小波滤波器的定义(见式(69a)、(69b)和(69c))．它的传递函数由

$$\begin{aligned}A(f)&\equiv\sum_{l=0}^{3}a_l e^{-i2\pi fl}\\&=-\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{2}}e^{-i2\pi f3}\\&=\frac{1}{\sqrt{2}}e^{-i3\pi f}\left[e^{i3\pi f}-e^{-i3\pi f}\right]\\&=\frac{1}{\sqrt{2}}e^{-i3\pi f}\left[-2i\sin(3\pi f)\right]\end{aligned}$$

给出．所以平方增益函数由$\mathcal{A}(f)\equiv|A(f)|^2=2\sin^2(3\pi f)$给出．图 519 给出了$\mathcal{A}(\cdot)$的图形，由图明显$\{a_l\}$不是高通滤波器．

练习[105b]答案 对滤波器$\{a_0=1,\ a_1=-1\}$的传递函数是

$$D(f)\equiv\sum_{l=0}^{1}a_l e^{-i2\pi fl}=1-e^{-i2\pi f}=e^{-i\pi f}[e^{i\pi f}-e^{-i\pi f}]=e^{-i\pi f}[2i\sin(\pi f)],$$

因此容易得到$\mathcal{D}(f)\equiv|D(f)|^2=4\sin^2(\pi f)$.

练习[106]答案 为了方便，令$M=L/2$,

$$C\equiv\cos^2(\pi f),\quad S\equiv\sin^2(\pi f),\quad M_l\equiv\binom{M+l}{l},\quad M_l^-\equiv\binom{M+l-1}{l},$$

从而

$$\mathcal{G}^{(D)}(f)=2C^M\sum_{l=0}^{M-1}M_l^-S^l\quad 和\quad \mathcal{H}^{(D)}(f)=2S^M\sum_{l=0}^{M-1}M_l^-C^l.$$

注意式(76b)，条件$\mathcal{H}^{(D)}(f)+\mathcal{H}^{(D)}\left(f+\frac{1}{2}\right)=2$等价于$\mathcal{G}^{(D)}(f)+\mathcal{H}^{(D)}(f)=2$，因此我们必须证明，对于所有的正整数$M$有

$$C^M\sum_{l=0}^{M-1}M_l^-S^l+S^M\sum_{l=0}^{M-1}M_l^-C^l=1. \tag{520a}$$

我们用归纳法证明．当$M=1$时直接得到，因为对于$M=1$，$M_0^-=1$并且$C+S=\cos^2(\pi f)+\sin^2(\pi f)=1$．下面是归纳过程．我们必须证明，给定式(520a)，由此得出

$$C^{M+1}\sum_{l=0}^{M}M_lS^l+S^{M+1}\sum_{l=0}^{M}M_lC^l=1. \tag{520b}$$

现在

$$\begin{aligned}C^{M+1}\sum_{l=0}^{M}M_lS^l&=C^M(1-S)\sum_{l=0}^{M}M_lS^l\\&=C^M\sum_{l=0}^{M}[M_lS^l-M_lS^{l+1}]\\&=C^M\left[\sum_{l=0}^{M}M_lS^l-\sum_{l=1}^{M+1}M_{l-1}S^l\right]\\&=C^M\left[1+\sum_{l=1}^{M}[M_l-M_{l-1}]S^l-M_MS^{M+1}\right].\end{aligned}$$

因为

$$\begin{aligned}M_l-M_{l-1}&=\binom{M+l}{l}-\binom{M+l-1}{l-1}\\&=\frac{(M+l)!}{M!l!}-\frac{(M+l-1)!}{M!(l-1)!}\\&=\frac{(M+l)!-(M+L-1)!\cdot l}{M!l!}\\&=\frac{(M+l-1)![M+l-l]}{M!l!}\\&=\frac{(M+l-1)!}{(M-1)!l!}\\&=\binom{M+l-1}{l}\\&=M_l^-,\end{aligned}$$

我们有

$$C^{M+1}\sum_{l=0}^{M}M_lS^l = C^M\Big[1+\sum_{l=1}^{M}M_{\bar{l}}S^l - M_MS^{M+1}\Big]$$

$$= C^M\Big[\sum_{l=0}^{M-1}M_{\bar{l}}S^l + M_{\bar{M}}S^M - M_MS^{M+1}\Big].$$

另外，式(520b)的第二项也可以表示为

$$S^{M+1}\sum_{l=0}^{M}M_lC^l = S^M\Big[\sum_{l=0}^{M-1}M_{\bar{l}}C^l + M_{\bar{M}}C^M - M_MC^{M+1}\Big].$$

因之，应用这两个表达式和式(520a)的归纳假设，产生

$$C^{M+1}\sum_{l=0}^{M}M_lS^l + S^{M+1}\sum_{l=0}^{M}M_lC^l$$

$$= 1 + C^M(M_{\bar{M}}S^M - M_MS^{M+1}) + S^M(M_{\bar{M}}C^M - M_MC^{M+1})$$

$$= 1 + 2C^MS^MM_{\bar{M}} - C^MS^MM_M(C+S)$$

$$= 1 + 2C^MS^MM_{\bar{M}} - C^MS^MM_M$$

$$= 1 + C^MS^M(2M_{\bar{M}} - M_M) = 1,$$

其中我们用到事实

$$2M_{\bar{M}} - M_M = 2\binom{2M-1}{M} - \binom{2M}{M}$$

$$= 2\frac{(2M-1)!}{M!(M-1)!} - \frac{(2M)!}{M!M!}$$

$$= \frac{2M\cdot(2M-1)! - (2M)!}{M!M!} = 0.$$

练习[111]答案 像要求的，我们有

$$U^{(\nu)}(f) = \sum_{l=-\infty}^{\infty}u_l^{(\nu)}e^{-i2\pi fl} = \sum_{l=-\infty}^{\infty}u_{l+\nu}e^{-i2\pi fl}$$

$$= \sum_{l=-\infty}^{\infty}u_l e^{-i2\pi f(l-\nu)}$$

$$= e^{i2\pi f\nu}\sum_{l=-\infty}^{\infty}u_l e^{-i2\pi fl} = e^{i2\pi f\nu}U(f).$$

练习[112]答案 式(76a)说明

$$G(f') = e^{-i2\pi f'(L-1)}H\Big(\frac{1}{2}-f'\Big),$$

因此我们有

$$H\Big(\frac{1}{2}-f'\Big) = e^{i2\pi f'(L-1)}G(f').$$

令 $f'=\frac{1}{2}-f$，使用 $e^{i\pi(L-1)}=e^{i\pi}$，因为 $L-1$ 总为奇数，我们得到

$$H(f) = e^{i2\pi\left(\frac{1}{2}-f\right)(L-1)}G\Big(\frac{1}{2}-f\Big)$$

$$= e^{-i2\pi f(L-1)+i\pi}G\Big(\frac{1}{2}-f\Big).$$

练习[114]答案 因为 $\theta^{(G)}(\cdot)$是对 $G(\cdot)$的相位函数，我们有 $G(f)=|G(f)|e^{i\theta^{(G)}(f)}$，因此

$$G_j(f)=\prod_{l=0}^{j-1}G(2^lf)=\prod_{l=0}^{j-1}[|G(2^lf)|e^{i\theta^{(G)}(2^lf)}]=\left[\prod_{l=0}^{j-1}|G(2^lf)|\right]e^{i\sum_{l=0}^{j-1}\theta^{(G)}(2^lf)},$$

则对 $G_j(\cdot)$的相位函数由

$$\theta_j^{(G)}(f)\equiv\sum_{l=0}^{j-1}\theta^{(G)}(2^lf)$$

给出．由于 $\sum_{l=0}^{j-1}2^l=2^j-1$，在上式中使用 $\theta^{(G)}(f)\approx2\pi f\nu$，则

$$\theta_j^{(G)}(f)\approx\sum_{l=0}^{j-1}2\pi(2^lf)\nu=2\pi f(2^j-1)\nu=2\pi f\nu_j^{(G)},\tag{522}$$

而 $\nu_j^{(G)}\equiv(2^j-1)\nu$．否则，因为 $\theta^{(H)}(\cdot)$是对 $H(\cdot)$的相位函数，则 $H(f)=|H(f)|e^{i\theta^{(H)}(f)}$．因此

$$H_j(f)=H(2^{j-1}f)G_{j-1}(f)=|H(2^{j-1}f)||G_{j-1}(f)|e^{i\theta^{(H)}(2^{j-1}f)}e^{i\theta_{j-1}^{(G)}(f)},$$

所以对 $H_j(\cdot)$的相位函数由

$$\theta_j^{(H)}(f)\equiv\theta^{(H)}(2^{j-1}f)+\theta_{j-1}^{(G)}(f)$$

给出．在上式应用 $\theta^{(H)}(f)\approx-2\pi f(L-1+\nu)$和式(522)得到

$$\theta_j^{(H)}(f)\approx-2\pi(2^{j-1}f)(L-1+\nu)+2\pi f(2^{j-1}-1)\nu=2\pi f[-2^{j-1}(L-1)-\nu]=2\pi f\nu_j^{(H)},$$

而 $\nu_j^{(H)}\equiv-[2^{j-1}(L-1)+\nu]$．

练习[116]答案 LA(8)尺度滤波器具有宽度 $L=8$，所以式(112e)表明 $\nu=-\frac{L}{2}+1=-3$．式(114b)现在表明

$$\nu_j^{(H)}=-(2^{j-1}[L-1]+\nu)=-7.2^{j-1}+3=\begin{cases}-4, & j=1;\\-11, & j=2;\\-25, & j=3;\end{cases}$$

另一方面，式(114a)表明

$$\nu_3^{(G)}=(2^3-1)\nu=-21.$$

由于这里 $\Delta t=1$，指标 t 和组成 $\mathbf{W}_1$ 的小波系数$\{W_{1,t}: t=0,1,\cdots,63\}$的实际时间之间结合，由表达式 $t_0+(2(t+1)-1+\nu_1^{(H)}\bmod N)$给出，所以我们有

$$\text{实际时间}=17+(2t-3\bmod 128)=\begin{cases}142, & t=0;\\144, & t=1;\\18, & t=2;\\20, & t=3;\\22, & t=4;\\\cdots\\138, & t=62;\\140, & t=63;\end{cases}$$

指标 t 和组成 $\mathbf{W}_2$ 的小波系数$\{W_{2,t}: t=0, 1, \cdots, 31\}$的实际时间之间的结合，由表达式 $t_0+(4(t+1)-1+\nu_2^{(H)} \bmod N)$给出，所以我们有

$$\text{实际时间} = 17+(4t-8 \bmod 128)=\begin{cases}137, & t=0;\\141, & t=1;\\17, & t=2;\\21, & t=3;\\25, & t=4;\\\cdots\\129, & t=30;\\133, & t=31;\end{cases}$$

指标 t 和组成 $\mathbf{W}_3$ 的小波系数$\{W_{3,t}: t=0, 1, \cdots, 15\}$的实际时间之间的结合，由表达式 $t_0+(8(t+1)-1+\nu_3^{(H)} \bmod N)$给出，所以，我们有

$$\text{实际时间} = 17+(8t-18 \bmod 128)=\begin{cases}127, & t=0;\\135, & t=1;\\143, & t=2;\\23, & t=3;\\31, & t=4;\\\cdots\\111, & t=14;\\119, & t=15;\end{cases}$$

最后，指标 t 和组成 $\mathbf{V}_3$ 的尺度系数$\{V_{3,t}: t=0, 1, \cdots, 15\}$的实际时间之间的结合，由表达式 $t_0+(8(t+1)-1+\nu_3^{(G)} \bmod N)$给出，所以，我们有

$$\text{实际时间} = 17+(8t-14 \bmod 128)=\begin{cases}131, & t=0;\\139, & t=1;\\19, & t=2;\\27, & t=3;\\35, & t=4;\\\cdots\\115, & t=14;\\123, & t=15.\end{cases}$$

图 524 给出了作为实际时间函数的 $\mathbf{W}_1$，$\mathbf{W}_2$，$\mathbf{W}_3$，$\mathbf{V}_3$ 元素的图.

练习[117]答案 用 $a_l=g_{L-1-l}$，$l=0, \cdots, L-1$ 定义反向滤波器. 那么

$$\begin{aligned}A(f) &\equiv \sum_{l=0}^{L-1} a_l e^{-i2\pi fl} = \sum_{l=0}^{L-1} g_{L-1-l} e^{-i2\pi fl} = \sum_{l=0}^{L-1} g_l e^{-i2\pi f(L-1-l)}\\&= e^{-i2\pi f(L-1)} \sum_{l=0}^{L-1} g_l e^{i2\pi fl}\\&= e^{-i2\pi f(L-1)} G^*(f),\end{aligned}$$

因为其中$\{g_l\}$是实值的.

练习[118]答案 由于 $\sum_l g_l^2 = \sum_l h_l^2 = 1$，我们有

$$e\{g_l\}+e\{h_l\} = \sum_{l=0}^{L-1} l g_l^2 + \sum_{l=0}^{L-1} l h_l^2.$$

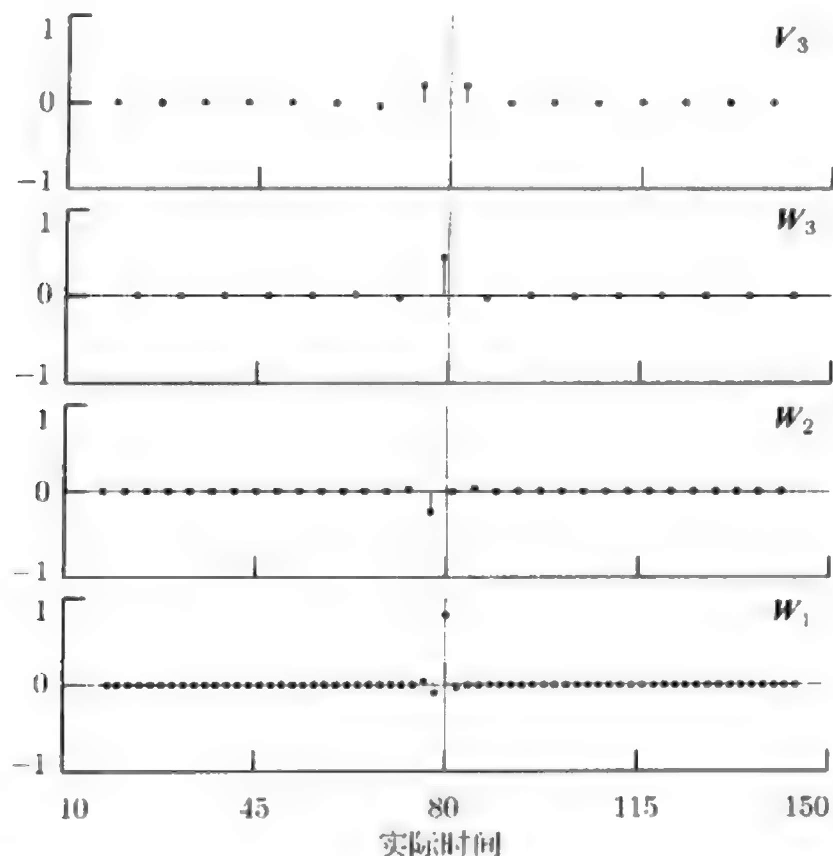


图 524 关于时间序列$\{X_t: t=0, \cdots, 127\}$的 $J_0=3$ 层的部分 LA(8)离散小波变换的小波和尺度系数，在 $X_{63}=1$ 时时间序列全为零. X_t 的第 t 个值与实际时间 $t+17$ 有关，因此 X_{63} 与实际时间 80 有关，在上图中用细垂线表示

应用式(75a)，则有

$$\sum_{l=0}^{L-1} l g_l^2 = \sum_{l=0}^{L-1} l h_{L-1-l}^2$$

$$= \sum_{l=0}^{L-1} (L-1-l) h_l^2,$$

从而导出所需要的

$$e\{g_l\} + e\{h_l\} = \sum_{l=0}^{L-1} (L-1-l) h_l^2 + \sum_{l=0}^{L-1} l h_l^2$$

$$= (L-1) \sum_{l=0}^{L-1} h_l^2 = L-1.$$

练习[139]答案 $\mathcal{D}_j$ 和 $\mathbf{W}_j$ 中的受影响项的部分分别为$(2^j L_j' + L_j - 2^j)/N$ 和 $L_j'/N_j = 2^j L_j'/N$. 对于大的 j，我们有 $L_j' = L-2$. 因此受影响的比率为

$$\frac{2^j L_j' + L_j - 2^j}{2^j L_j'} = \frac{2^j(L-2) + (2^j-1)(L-1) + 1 - 2^j}{2^j(L-2)}$$

$$= \frac{2^{j+1}-1}{2^j} \approx 2.$$

练习[149]答案 由于 $L_j'>0$ 当且仅当 $L\geqslant 4$，我们有

$$L_j = (2^j-1)(L-1)+1 \geqslant 3\cdot 2^j - 2 > 2^j.$$

为了证明 $L_j>2^jL_j'$，令 $L_j'=r+\delta$，而 $r=(L-2)\left(1-\frac{1}{2^j}\right)$ 和 $0\leqslant\delta<1$，那么

$$\begin{aligned}2^jL_j' &= 2^jr+2^j\delta = (L-2)(2^j-1)+2^j\delta \\ &= (2^j-1)(L-1)+1-(2^j-1)-1+2^j\delta = L_j+2^j(\delta-1).\end{aligned}$$

因此

$$L_j = 2^jL'_j+2^j(1-\delta) > 2^jL'_j.$$

练习[166]答案 注意到

$$\frac{1}{2}\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1 = \frac{1}{2}\begin{bmatrix} h_1 & h_3 & \cdots \\ h_0 & h_2 & \cdots \\ 0 & h_1 & \cdots \\ 0 & h_0 & \cdots \\ \vdots & \vdots & \cdots \end{bmatrix}\begin{bmatrix} \sqrt{2}\,\widetilde{W}_{1,1} \\ \sqrt{2}\,\widetilde{W}_{1,3} \\ \vdots \end{bmatrix} = \begin{bmatrix} \tilde{h}_1 & \tilde{h}_3 & \cdots \\ \tilde{h}_0 & \tilde{h}_2 & \cdots \\ 0 & \tilde{h}_1 & \cdots \\ 0 & \tilde{h}_0 & \cdots \\ \vdots & \vdots & \cdots \end{bmatrix}\begin{bmatrix} \widetilde{W}_{1,1} \\ \widetilde{W}_{1,3} \\ \vdots \end{bmatrix}$$

和

$$\frac{1}{2}\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1} = \frac{1}{2}\begin{bmatrix} h_0 & h_2 & \cdots \\ 0 & h_1 & \cdots \\ 0 & h_0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \cdots \end{bmatrix}\begin{bmatrix} \sqrt{2}\,\widetilde{W}_{1,0} \\ \sqrt{2}\,\widetilde{W}_{1,2} \\ \vdots \end{bmatrix} = \begin{bmatrix} \tilde{h}_0 & \tilde{h}_2 & \cdots \\ 0 & \tilde{h}_1 & \cdots \\ 0 & \tilde{h}_0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \cdots \end{bmatrix}\begin{bmatrix} \widetilde{W}_{1,0} \\ \widetilde{W}_{1,2} \\ \vdots \end{bmatrix}.$$

交叉存取上面两式，我们能够写

$$\frac{1}{2}\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1+\frac{1}{2}\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1} = \begin{bmatrix} \tilde{h}_0 & \tilde{h}_1 & \tilde{h}_2 & \tilde{h}_3 & \cdots \\ 0 & \tilde{h}_0 & \tilde{h}_1 & \tilde{h}_2 & \cdots \\ 0 & 0 & \tilde{h}_0 & \tilde{h}_1 & \cdots \\ 0 & 0 & 0 & \tilde{h}_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix}\begin{bmatrix} \widetilde{W}_{1,0} \\ \widetilde{W}_{1,1} \\ \widetilde{W}_{1,2} \\ \widetilde{W}_{1,3} \\ \vdots \end{bmatrix} = \widetilde{\mathcal{B}}_1\widetilde{\mathbf{W}}_1.$$

类似的论证产生$\frac{1}{2}\mathcal{A}_1^{\mathrm{T}}\mathbf{V}_1+\frac{1}{2}\mathcal{A}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{V}_{\mathcal{T},1}=\widetilde{\mathcal{A}}_1\widetilde{\mathbf{V}}_1$.

练习[167a]答案 由于$\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\mathbf{W}}_1=\frac{1}{2}(\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1+\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1})$并且由于$\mathcal{B}_1\mathcal{B}_1^{\mathrm{T}}=I_{N/2}$及 $\mathcal{B}_{\mathcal{T},1}\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}=I_{N/2}$，则有

$$\begin{aligned}\|\widetilde{\mathcal{D}}_1\|^2 &= \|\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\mathbf{W}}_1\|^2 = \frac{1}{4}\|\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1+\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}\|^2 \\ &= \frac{1}{4}(\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1+\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1})^{\mathrm{T}}(\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1+\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}) \\ &= \frac{1}{4}(\mathbf{W}_1^{\mathrm{T}}\mathcal{B}_1\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1+\mathbf{W}_{\mathcal{T},1}^{\mathrm{T}}\mathcal{B}_{\mathcal{T},1}\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}+\mathbf{W}_1^{\mathrm{T}}\mathcal{B}_1\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}+\mathbf{W}_{\mathcal{T},1}^{\mathrm{T}}\mathcal{B}_{\mathcal{T},1}\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1) \\ &= \frac{1}{4}(\mathbf{W}_1^{\mathrm{T}}\mathbf{W}_1+\mathbf{W}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}+\mathbf{W}_1^{\mathrm{T}}\mathcal{B}_1\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}+\mathbf{W}_{\mathcal{T},1}^{\mathrm{T}}\mathcal{B}_{\mathcal{T},1}\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1) \\ &= \frac{1}{4}(\|\mathbf{W}_1\|^2+\|\mathbf{W}_{\mathcal{T},1}\|^2+\mathbf{W}_1^{\mathrm{T}}\mathcal{B}_1\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}+\mathbf{W}_{\mathcal{T},1}^{\mathrm{T}}\mathcal{B}_{\mathcal{T},1}\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1).\end{aligned}$$

由于 $\|\widetilde{\mathbf{W}}_1\|^2=\frac{1}{2}(\|\mathbf{W}_1\|^2+\|\mathbf{W}_{\mathcal{T},1}\|^2)$，并且标量 $\mathbf{W}_1^{\mathrm{T}}\mathcal{B}_1\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}$ 等于它的转置 $\mathbf{W}_{\mathcal{T},1}^{\mathrm{T}}\mathcal{B}_{\mathcal{T},1}\mathcal{B}_1^{\mathrm{T}}\mathbf{W}_1$，我们得到

$$\|\widetilde{\mathcal{D}}_1\|^2 = \frac{1}{2}(\|\widetilde{\mathbf{W}}_1\|^2+\mathbf{W}_1^{\mathrm{T}}\mathcal{B}_1\mathcal{B}_{\mathcal{T},1}^{\mathrm{T}}\mathbf{W}_{\mathcal{T},1}).$$

练习[167b]答案 令$\{\tilde h_l^\circ\}$是$\{\tilde h_l\}$的长为 N 的周期化序列：

$$\tilde h_l^\circ \equiv \begin{cases} \tilde h_l, & 0 \leqslant l \leqslant L-1; \\ 0, & L \leqslant l \leqslant N-1; \\ \tilde h^\circ_{l \bmod N}, & l < 0 \text{ or } l \geqslant N. \end{cases}$$

以类似的方法定义$\{\tilde g_l^\circ\}$. 那么$\widetilde{\mathcal{B}}_1^{\mathrm{T}}\,\widetilde{\mathcal{B}}_1$ 的第$(j,\ k)$个元素由

$$b_{j,k} = \sum_{l=0}^{N-1} \tilde h^\circ_{j+l} \tilde h^\circ_{k+l} = \sum_{l=0}^{N-1} \tilde h^\circ_l \tilde h^\circ_{k-j+l}$$

$$= \sum_{l=0}^{L-1} \tilde h^\circ_l \tilde h^\circ_{k-j+l}, \quad 0 \leqslant j,k \leqslant N-1$$

给出，并且同样地存在$\widetilde{\mathcal{A}}_1^{\mathrm{T}}\,\widetilde{\mathcal{A}}_1$ 的第$(j,\ k)$个元素 $a_{j,k}$ 的类似表示. 因此$\widetilde{\mathcal{B}}_1^{\mathrm{T}}\,\widetilde{\mathcal{B}}_1 + \widetilde{\mathcal{A}}_1^{\mathrm{T}}\,\widetilde{\mathcal{A}}_1$ 的第$(j,\ k)$个元素由 $b_{j,k}+a_{j,k}$ 给出，所以只需证明在 $j=k$ 时 $b_{j,k}+a_{j,k}=1$，否则等于 0. 当 $j=k$ 时，我们从式(163a)的中间部分得到

$$b_{k,k} = \sum_{l=0}^{L-1} [\tilde h^\circ_l]^2 = \sum_{l=0}^{L-1} \tilde h_l^2 = \frac{1}{2},$$

我们类似地有 $a_{k,k}=\frac{1}{2}$，由此得到 $b_{k,k}+a_{k,k}=1$. 如果对于一些非零整数 m 满足 $k-j=2m$，那么

$$b_{j,k} = \sum_{l=0}^{L-1} \tilde h^\circ_l \tilde h^\circ_{2m+l} = \sum_{l=0}^{L-1} \tilde h_l\, \tilde h_{2m+l \bmod N} = 0.$$

因为上面第二个和对于一些 m'可以写成一个或者两个的形式

$$\sum_{l=0}^{L-1} \tilde h_l\, \tilde h_{2m'+l}$$

的和，由式(163 a)的最后部分它等于零. 对于 $a_{j,k}$ 有类似的结论，所以当 $|k-j|$ 是正偶数时，$b_{j,k}+a_{j,k}=0$. 最后，如果 $k-j$ 是奇数，首先注意式(163d)的第一个关系式和$\{\tilde g_l^\circ\}$以及$\{\tilde h_l^\circ\}$的定义告诉我们

$$g_l^\circ = (-1)^{l+1}\, \tilde h^\circ_{L-l-1} \qquad \text{对于所有的 } l.$$

因此我们有

$$b_{j,k}+a_{j,k} = \sum_{l=0}^{N-1} \tilde h_l\, \tilde h_{k-j+l} + \sum_{l=0}^{N-1} \tilde g_l\, \tilde g_{k-j+l}$$

$$= \sum_{l=0}^{N-1} \tilde h_l\, \tilde h_{k-j+l} + \sum_{l=0}^{N-1} (-1)^{l+1}\, \tilde h_{L-l-1} (-1)^{k-j+l+1}\, \tilde h_{L-(k-j+l)-1}$$

$$= \sum_{l=0}^{N-1} \tilde h_l\, \tilde h_{k-j+l} - \sum_{l=0}^{N-1} \tilde h_{L-l-1}\, \tilde h_{L-k+j-l-1}$$

$$= \sum_{l=0}^{N-1} \tilde h_l\, \tilde h_{k-j+l} - \sum_{l=0}^{N-1} \tilde h_{-l}\, \tilde h_{-k+j-l}$$

$$= \sum_{l=0}^{N-1} \tilde h_l\, \tilde h_{k-j+l} - \sum_{l=0}^{N-1} \tilde h_l\, \tilde h_{l-k+j}$$

$$= \sum_{l=0}^{N-1} \tilde h_l\, \tilde h_{k-j+l} - \sum_{l=0}^{N-1} \tilde h_{l+k-j}\, \tilde h_l = 0,$$

这就证明了提示.

由式(165b)，我们有$\widetilde{\boldsymbol{W}}_1=\widetilde{\mathcal{B}}_1\boldsymbol{X}$和$\widetilde{\boldsymbol{V}}_1=\widetilde{\mathcal{A}}_1\boldsymbol{X}$. 分别用$\widetilde{\mathcal{B}}_1^{\mathrm{T}}$左乘第一个式，用$\widetilde{\mathcal{A}}_1^{\mathrm{T}}$左乘第二个式两边，然后相加得到

$$\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\boldsymbol{W}}_1+\widetilde{\mathcal{A}}_1^{\mathrm{T}}\widetilde{\boldsymbol{V}}_1=\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\mathcal{B}}_1\boldsymbol{X}+\widetilde{\mathcal{A}}_1^{\mathrm{T}}\widetilde{\mathcal{A}}_1\boldsymbol{X}=(\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\mathcal{B}}_1+\widetilde{\mathcal{A}}_1^{\mathrm{T}}\widetilde{\mathcal{A}}_1)\boldsymbol{X}=\boldsymbol{X},$$

其中我们使用了提示. 用$\boldsymbol{X}^{\mathrm{T}}$左乘上式的两边得到

$$\boldsymbol{X}^{\mathrm{T}}\widetilde{\mathcal{B}}_1^{\mathrm{T}}\widetilde{\boldsymbol{W}}_1+\boldsymbol{X}^{\mathrm{T}}\widetilde{\mathcal{A}}_1^{\mathrm{T}}\widetilde{\boldsymbol{V}}_1=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X};$$

然而，由于$\widetilde{\boldsymbol{W}}_1^{\mathrm{T}}=\boldsymbol{X}^{\mathrm{T}}\widetilde{\mathcal{B}}_1^{\mathrm{T}}$和$\widetilde{\boldsymbol{V}}_1^{\mathrm{T}}=\boldsymbol{X}^{\mathrm{T}}\widetilde{\mathcal{A}}_1^{\mathrm{T}}$，上式变成

$$\widetilde{\boldsymbol{W}}_1^{\mathrm{T}}\widetilde{\boldsymbol{W}}_1+\widetilde{\boldsymbol{V}}_1^{\mathrm{T}}\widetilde{\boldsymbol{V}}_1=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X},$$

即得所求

$$\|\boldsymbol{W}_1\|^2+\|\boldsymbol{V}_1\|^2=\|\boldsymbol{X}\|^2.$$

练习[171a]答案 令$\{\tilde{h}_l^\circ\}$和$\{\tilde{g}_l^\circ\}$是周期化$\{\tilde{h}_l\}=\{h_l/\sqrt{2}\}$和$\{\tilde{g}_l\}=\{g_l/\sqrt{2}\}$获得的长度为$N$的滤波器. 注意到$\{\tilde{h}_l^\circ\}\leftrightarrow\left\{\widetilde{H}\left(\frac{k}{N}\right)\right\}$和$\{\tilde{g}_l^\circ\}\leftrightarrow\left\{\widetilde{G}\left(\frac{k}{N}\right)\right\}$. 现在，

$$\widetilde{W}_{1,t}=\sum_{l=0}^{N-1}\tilde{h}_l^\circ X_{t-l \bmod N}\quad 和 \quad \widetilde{V}_{1,t}=\sum_{l=0}^{N-1}\tilde{g}_l^\circ X_{t-l \bmod N}.$$

令$\{\mathcal{X}_k\}$为$\{X_t\}$的离散傅里叶变换，则有

$$\{\widetilde{W}_{1,t}\}\leftrightarrow\left\{\widetilde{H}\left(\frac{k}{N}\right)\mathcal{X}_k\right\}\quad 和 \quad \{\widetilde{V}_{1,t}\}\leftrightarrow\left\{\widetilde{G}\left(\frac{k}{N}\right)\mathcal{X}_k\right\}.$$

用帕塞瓦尔定理(式(36h))可得

$$\|\widetilde{\boldsymbol{W}}_1\|^2=\frac{1}{N}\sum_{k=0}^{N-1}\left|\widetilde{H}\left(\frac{k}{N}\right)\right|^2|\mathcal{X}_k|^2=\frac{1}{N}\sum_{k=0}^{N-1}\widetilde{\mathcal{H}}\left(\frac{k}{N}\right)|\mathcal{X}_k|^2$$

和

$$\|\widetilde{\boldsymbol{V}}_1\|^2=\frac{1}{N}\sum_{k=0}^{N-1}\left|\widetilde{G}\left(\frac{k}{N}\right)\right|^2|\mathcal{X}_k|^2=\frac{1}{N}\sum_{k=0}^{N-1}\widetilde{\mathcal{G}}\left(\frac{k}{N}\right)|\mathcal{X}_k|^2.$$

相加可得

$$\|\widetilde{\boldsymbol{W}}_1\|^2+\|\widetilde{\boldsymbol{V}}_1\|^2=\frac{1}{N}\sum_{k=0}^{N-1}|\mathcal{X}_k|^2\left(\widetilde{\mathcal{H}}\left(\frac{k}{N}\right)+\widetilde{\mathcal{G}}\left(\frac{k}{N}\right)\right)=\frac{1}{N}\sum_{k=0}^{N-1}|\mathcal{X}_k|^2=\|\boldsymbol{X}\|^2,$$

其中我们应用了(ⅰ)对于所有的f，$\widetilde{\mathcal{H}}(f)+\widetilde{\mathcal{G}}(f)=1$(参见式(163c))和(ⅱ)再次用到帕塞瓦尔定理.

练习[171b]答案 $\widetilde{\boldsymbol{V}}_{J_0}$中的元素只有$\{\widetilde{V}_{J_0,t}\}$，这是用$\{\tilde{g}_{J_0,l}^\circ\}$循环卷积$\{X_t\}$的结果(参见式(170a)). 由于$\{\tilde{g}_{J_0,l}^\circ\}$是周期化$\{\tilde{g}_{J_0,l}\}$为长$N$的结果，并且$\{\tilde{g}_{J_0,l}\}\leftrightarrow\widetilde{G}_{J_0}(\cdot)$，我们有$\{\tilde{g}_{J_0,l}^\circ\}\leftrightarrow\left\{\widetilde{G}_{J_0}\left(\frac{k}{N}\right)\right\}$(这由练习[33]得到). 如果让$\{\mathcal{X}_k\}$为$\{X_t\}$的离散傅里叶变换，则有

$$\{\widetilde{V}_{J_0,t}\}\leftrightarrow\left\{\widetilde{G}_{J_0}\left(\frac{k}{N}\right)\mathcal{X}_k\right\}.$$

其次注意，如果$\{a_t: t=0,\cdots,N-1\}\leftrightarrow\{A_k: k=0,\cdots,N-1\}$，那么$A_0=N\bar{a}$，其中$\bar{a}$是$\{a_t\}$的样本均值. 现在得到$\{\widetilde{V}_{J_0,t}\}$的样本均值由

$$\frac{1}{N}\widetilde{G}_{J_0}(0)\mathcal{X}_0=\frac{1}{N}\left(\prod_{l=0}^{J_0-1}\widetilde{G}(0)\right)N\overline{X}=\overline{X}$$

给出，从练习[76b]中知$G(0)=\sum_l g_l=\sqrt{2}$，所以$\widetilde{G}(0)=G(0)/\sqrt{2}=1$. 在关系式$\hat{\sigma}_X^2=\frac{1}{N}\|\boldsymbol{X}\|^2-\overline{X}^2$中用$\{\widetilde{V}_{J_0,t}\}$替换$\{X_t\}$，现在得到$\{\widetilde{V}_{J_0,t}\}$的样本方差$\frac{1}{N}\|\widetilde{\boldsymbol{V}}_{J_0}\|^2-\overline{X}^2$.

练习[172]答案 $\widetilde{\mathcal{D}}_1$ 和$\widetilde{\mathcal{S}}_1$ 中的元素由

$$\widetilde{\mathcal{D}}_{1,t}=\sum_{l=0}^{N-1}\tilde{h}_l^{\circ}\widetilde{W}_{1,t+l \bmod N} \quad 和 \quad \widetilde{\mathcal{S}}_{1,t}=\sum_{l=0}^{N-1}\tilde{g}_l^{\circ}\widetilde{V}_{1,t+l \bmod N}$$

给出，$t=0,1,\cdots,N-1$. 如前，$\{\widetilde{\mathcal{D}}_{1,t}\}$可以认为是循环滤波$\{\widetilde{W}_{1,t}\}$，滤波器的离散傅里叶变换是$\left\{\widetilde{H}^{*}\left(\frac{k}{N}\right)\right\}$；另一方面，$\{\widetilde{W}_{1,t}\}$可以认为是具有滤波器$\{h_l^{\circ}\}$的循环滤波$\{X_t\}$滤波器$\{\tilde{h}_l^{\circ}\}$的离散傅里叶变换$\left\{\widetilde{H}\left(\frac{k}{N}\right)\right\}$. 令$\{\mathcal{X}_k\}$为$\{X_t\}$的离散傅里叶变换，则有

$$\{\widetilde{\mathcal{D}}_{1,t}\}\leftrightarrow\left\{\widetilde{H}^{*}\left(\frac{k}{N}\right)\widetilde{H}\left(\frac{k}{N}\right)\mathcal{X}_k\right\}=\left\{\left|\widetilde{H}\left(\frac{k}{N}\right)\right|^2\mathcal{X}_k\right\}=\left\{\widetilde{\mathcal{H}}\left(\frac{k}{N}\right)\mathcal{X}_k\right\},$$

同样$\{\widetilde{\mathcal{S}}_{1,t}\}\leftrightarrow\left\{\mathcal{G}\left(\frac{k}{N}\right)\mathcal{X}_k\right\}$. 相加可得

$$\{\widetilde{\mathcal{D}}_{1,t}+\widetilde{\mathcal{S}}_{1,t}\}\leftrightarrow\left\{\left(\widetilde{\mathcal{H}}\left(\frac{k}{N}\right)+\widetilde{\mathcal{G}}\left(\frac{k}{N}\right)\right)\mathcal{X}_k\right\}=\{\mathcal{X}_k\},$$

由于对于所有的f有$\mathcal{H}(f)+\mathcal{G}(f)=1$(参见式(163b)). 由于$\{X_t\}\leftrightarrow\{\mathcal{X}_k\}$，对于所有的$t$我们一定有$X_t=\widetilde{\mathcal{D}}_{1,t}+\widetilde{\mathcal{S}}_{1,t}$，即$\mathbf{X}=\widetilde{\mathcal{D}}_1+\widetilde{\mathcal{S}}_1$.

练习[175]答案 令$\{\hat{W}_{j,k}\}$是$\{\widetilde{W}_{j,t}\}$的离散傅里叶变换，令$\{\hat{V}_{j,k}\}$为$\{\widetilde{V}_{j,t}\}$的离散傅里叶变换. 由式(175b)告诉我们$\{\widetilde{W}_{j,t}\}$由循环滤波$\{\widetilde{V}_{j-1,t}\}$形成，其滤波器的传递函数由$\widetilde{H}(2^{j-1}f)$给出，则有

$$\hat{W}_{j,k}=\widetilde{H}\left(2^{j-1}\frac{k}{N}\right)\hat{V}_{j-1,k};$$

同样，从式(175c)得

$$\hat{V}_{j,k}=\widetilde{G}\left(2^{j-1}\frac{k}{N}\right)\hat{V}_{j-1,k}.$$

式(175d)的右边第一个和是与$\{\widetilde{W}_{j,t}\}$滤波器的互相关，而滤波器的传递函数由$\widetilde{H}(2^{j-1}f)$定义，所以，我们有

$$\left\{\sum_{l=0}^{L-1}\tilde{h}_l\widetilde{W}_{j,t+2^{j-1}l \bmod N}\right\}\leftrightarrow\left\{\widetilde{H}^{*}\left(2^{j-1}\frac{k}{N}\right)\hat{W}_{j,k}\right\};$$

同样，对于第二个和有

$$\left\{\sum_{l=0}^{L-1}\tilde{g}_l\widetilde{V}_{j,t+2^{j-1}l \bmod N}\right\}\leftrightarrow\left\{\widetilde{G}^{*}\left(2^{j-1}\frac{k}{N}\right)\hat{V}_{j,k}\right\}.$$

式(175d)右边的离散傅里叶变换因此由

$$\begin{aligned}
&\widetilde{H}^{*}\left(2^{j-1}\frac{k}{N}\right)\hat{W}_{j,k}+\widetilde{G}^{*}\left(2^{j-1}\frac{k}{N}\right)\hat{V}_{j,k}\\
&\quad=\widetilde{H}^{*}\left(2^{j-1}\frac{k}{N}\right)\widetilde{H}\left(2^{j-1}\frac{k}{N}\right)\hat{V}_{j-1,k}+\widetilde{G}^{*}\left(2^{j-1}\frac{k}{N}\right)\widetilde{G}\left(2^{j-1}\frac{k}{N}\right)\hat{V}_{j-1,k}\\
&\quad=\left(\left|\widetilde{H}\left(2^{j-1}\frac{k}{N}\right)\right|^2+\left|\widetilde{G}\left(2^{j-1}\frac{k}{N}\right)\right|^2\right)\hat{V}_{j-1,k}\\
&\quad=\hat{V}_{j-1,k}
\end{aligned}$$

给出，由于对于所有的f有$|\widetilde{H}(f)|^2+|\widetilde{G}(f)|^2=\widetilde{\mathcal{H}}(f)+\widetilde{\mathcal{G}}(f)=1$(参见式(163b))；即式(175d)的右边等于$\widetilde{V}_{j-1,t}$.

练习[208]答案 我们有

$$
\begin{bmatrix} \mathcal{A}_2\mathcal{B}_1 \\ \mathcal{B}_2\mathcal{B}_1 \\ \mathcal{A}_1 \end{bmatrix}
\begin{bmatrix} \mathcal{B}_1^{\mathrm{T}}\mathcal{A}_2^{\mathrm{T}} & \mathcal{B}_1^{\mathrm{T}}\mathcal{B}_2^{\mathrm{T}} & \mathcal{A}_1^{\mathrm{T}} \end{bmatrix}
= \begin{bmatrix} \mathcal{A}_2\mathcal{B}_1\mathcal{B}_1^{\mathrm{T}}\mathcal{A}_2^{\mathrm{T}} & \mathcal{A}_2\mathcal{B}_1\mathcal{B}_1^{\mathrm{T}}\mathcal{B}_2^{\mathrm{T}} & \mathcal{A}_2\mathcal{B}_1\mathcal{A}_1^{\mathrm{T}} \\ \mathcal{B}_2\mathcal{B}_1\mathcal{B}_1^{\mathrm{T}}\mathcal{A}_2^{\mathrm{T}} & \mathcal{B}_2\mathcal{B}_1\mathcal{B}_1^{\mathrm{T}}\mathcal{B}_2^{\mathrm{T}} & \mathcal{B}_2\mathcal{B}_1\mathcal{A}_1^{\mathrm{T}} \\ \mathcal{A}_1\mathcal{B}_1^{\mathrm{T}}\mathcal{A}_2^{\mathrm{T}} & \mathcal{A}_1\mathcal{B}_1^{\mathrm{T}}\mathcal{B}_2^{\mathrm{T}} & \mathcal{A}_1\mathcal{A}_1^{\mathrm{T}} \end{bmatrix}
$$

$$
= \begin{bmatrix} \mathcal{A}_2 I_{\frac{N}{2}}\mathcal{A}_2^{\mathrm{T}} & \mathcal{A}_2 I_{\frac{N}{2}}\mathcal{B}_2^{\mathrm{T}} & \mathcal{A}_2 0_{\frac{N}{2}} \\ \mathcal{B}_2 I_{\frac{N}{2}}\mathcal{A}_2^{\mathrm{T}} & \mathcal{B}_2 I_{\frac{N}{2}}\mathcal{B}_2^{\mathrm{T}} & \mathcal{B}_2 0_{\frac{N}{2}} \\ 0_{\frac{N}{2}}\mathcal{A}_2^{\mathrm{T}} & 0_{\frac{N}{2}}\mathcal{B}_2^{\mathrm{T}} & I_{\frac{N}{2}} \end{bmatrix}
$$

$$
= \begin{bmatrix} \mathcal{A}_2\mathcal{A}_2^{\mathrm{T}} & \mathcal{A}_2\mathcal{B}_2^{\mathrm{T}} & 0_{\frac{N}{4},\frac{N}{2}} \\ \mathcal{B}_2\mathcal{A}_2^{\mathrm{T}} & \mathcal{B}_2\mathcal{B}_2^{\mathrm{T}} & 0_{\frac{N}{4},\frac{N}{2}} \\ 0_{\frac{N}{2},\frac{N}{4}} & 0_{\frac{N}{2},\frac{N}{4}} & I_{\frac{N}{2}} \end{bmatrix}
$$

$$
= \begin{bmatrix} I_{\frac{N}{4}} & 0_{\frac{N}{4}} & 0_{\frac{N}{4},\frac{N}{2}} \\ 0_{\frac{N}{4}} & I_{\frac{N}{4}} & 0_{\frac{N}{4},\frac{N}{2}} \\ 0_{\frac{N}{2},\frac{N}{4}} & 0_{\frac{N}{2},\frac{N}{4}} & I_{\frac{N}{2}} \end{bmatrix} = I_N.
$$

练习[214]答案 首先注意(从式(214a)开始)

$$
\boldsymbol{W}_{j,2n,t} = \sum_{l=0}^{L-1} u_{2n,l}\boldsymbol{W}_{j-1,n,2t+1-l \bmod N_j-1},
$$

而

$$
\boldsymbol{W}_{j,2n+1,t} = \sum_{l=0}^{L-1} u_{2n+1,l}\boldsymbol{W}_{j-1,n,2t+1-l \bmod N_j-1}.
$$

为建立练习结论，我们只需证明 $u_{2n,l}=a_{n,l}$ 和 $u_{2n+1,l}=b_{n,l}$. 在上面第一式中，$u_{2n,l}$ 的第一个下标是偶数，所以

$$
u_{2n,l} = \begin{cases} g_l, & 2n \bmod 4 = 0,\text{即 } n \text{ 是偶数}; \\ h_l, & 2n \bmod 4 = 2,\text{即 } n \text{ 是奇数}. \end{cases}
$$

因之，$u_{2n,l}=a_{n,l}$. 在第二式中，$u_{2n+1,l}$ 第一个下标是奇数，所以

$$
u_{2n+1,l} = \begin{cases} h_l, & 2n+1 \bmod 4 = 1,\text{即 } n \text{ 是偶数}; \\ g_l, & 2n+1 \bmod 4 = 3,\text{即 } n \text{ 是奇数}. \end{cases}
$$

因之，$u_{2n+1,l}=b_{n,l}$.

练习[223]答案 如果我们使用 $m(|W_{j,n,t}|)=|W_{j,n,t}|^2$，那么对于任何的 $\mathcal{C}\subset\mathcal{N}$，我们都有

$$
\sum_{(j,n)\in\mathcal{C}} M(\boldsymbol{W}_{j,n}) = \sum_{(j,n)\in\mathcal{C}} \|\boldsymbol{W}_{j,n}\|^2 = \|\boldsymbol{X}\|^2,
$$

因为每个变换都是规范的. 因此，对于这个价值泛函，每一个可能的变换都要满足最优性准则.

练习[230]答案 由于对于所有的最接近对称和 coiflet 滤波器，$\nu<0$，我们能够写

$$\nu_{j,n}=-|\nu|(2^j-1)-S_{j,n,1}(L-1-2|\nu|),$$

所以 $\nu_{j,n}<0$ 当且仅当

$$S_{j,n,1}(2|\nu|+1-L)<|\nu|(2^j-1).$$

由于 $|\nu|(2^j-1)>0$ 和 $2^j-1\geqslant S_{j,n,1}\geqslant 0$，上面的不等式成立当且仅当$(2|\nu|+1-L)<|\nu|$，即 $|\nu|<L-1$. 从式(112e)我们可以推出对于所有的最接近对称滤波器，$|\nu|\leqslant\frac{L}{2}-2$；由于$\frac{L}{2}-2<L-1$ 总是真的，由此得出 $|\nu|<L-1$ 成立. 对于 coiflet 滤波器，我们总有 $|\nu|=\frac{2L}{3}-1<L-1$，所以结论成立.

练习[233]答案 如果 $j=1$，我们有

$$\begin{aligned}\sum_{n=0}^{1}|\widetilde{U}_{1,n}(f)|^2&=|\widetilde{M}_{c_{1,0,0}}(f)|^2+|\widetilde{M}_{c_{1,1,0}}(f)|^2\\&=|\widetilde{M}_0(f)|^2+|\widetilde{M}_1(f)|^2\\&=|\widetilde{G}_0(f)|^2+|\widetilde{H}_1(f)|^2=1,\end{aligned}$$

其中我们回忆 $\mathbf{c}_{1,0}\equiv[0]=[c_{1,0,0}]$和 $\mathbf{c}_{1,1}\equiv[1]=[c_{1,1,0}]$，并且上式中最后一个等式就是式(163a). 因此这个练习在 $j=1$时成立. 为看到对 $j>1$ 时也成立. 我们要求

$$\sum_{n=0}^{2^j-1}|\widetilde{U}_{j,n}(f)|^2=\sum_{n=0}^{2^{j-1}-1}|\widetilde{U}_{j-1,n}(f)|^2,$$

我们准备由归纳法建立练习的结论. 为了证明上式成立，注意

$$\sum_{n=0}^{2^j-1}|\widetilde{U}_{j,n}(f)|^2=\sum_{n=0}^{2^j-1}\prod_{m=0}^{j-1}|\widetilde{M}_{c_{j,n,m}}(2^mf)|^2, \tag{530}$$

并且考虑前两项相加：

$$\begin{aligned}&\prod_{m=0}^{j-1}|\widetilde{M}_{c_{j,0,m}}(2^mf)|^2+\prod_{m=0}^{j-1}|\widetilde{M}_{c_{j,1,m}}(2^mf)|^2\\&=\Big(\prod_{m=0}^{j-2}|\widetilde{M}_{c_{j,0,m}}(2^mf)|^2\Big)|\widetilde{M}_{c_{j,0,j-1}}(2^mf)|^2\\&\quad+\Big(\prod_{m=0}^{j-2}|\widetilde{M}_{c_{j,1,m}}(2^mf)|^2\Big)|\widetilde{M}_{c_{j,1,j-1}}(2^mf)|^2\\&=\Big(\prod_{m=0}^{j-2}|\widetilde{M}_{c_{j-1,0,m}}(2^mf)|^2\Big)\big(|\widetilde{M}_0(2^mf)|^2+|\widetilde{M}_1(2^mf)|^2\big)\\&=\Big(\prod_{m=0}^{j-2}|\widetilde{M}_{c_{j-1,0,m}}(2^mf)|^2\Big)\big(|\widetilde{G}(2^mf)|^2+|\widetilde{H}(2^mf)|^2\big)\\&=\prod_{m=0}^{j-2}|\widetilde{M}_{c_{j-1,0,m}}(2^mf)|^2=|\widetilde{U}_{j-1,0}(f)|^2;\end{aligned}$$

这里我们应用事实：$\boldsymbol{c}_{j,0}$和$\boldsymbol{c}_{j,1}$的前$j-1$个元素和$\boldsymbol{c}_{j-1,0}$中的元素是相同的，而最后的元素分别是0和1. 以类似的方法，式(530)中接下来的两项产生

$$
\begin{aligned}
&\prod_{m=0}^{j-1}|\widetilde{M}_{c_j,2,m}(2^m f)|^2+\prod_{m=0}^{j-1}|\widetilde{M}_{c_j,3,m}(2^m f)|^2\\
&=\Big(\prod_{m=0}^{j-2}|\widetilde{M}_{c_{j-1},1,m}(2^m f)|^2\Big)\big(|\widetilde{M}_1(2^m f)|^2+|\widetilde{M}_0(2^m f)|^2\big)\\
&=\Big(\prod_{m=0}^{j-2}|\widetilde{M}_{c_{j-1},1,m}(2^m f)|^2\Big)\big(|\widetilde{H}(2^m f)|^2+|\widetilde{G}(2^m f)|^2\big)\\
&=\prod_{m=0}^{j-2}|\widetilde{M}_{c_{j-1},1,m}(2^m f)|^2=|\widetilde{U}_{j-1,1}(f)|^2;
\end{aligned}
$$

其中现在我们回忆：$\boldsymbol{c}_{j,2}$和$\boldsymbol{c}_{j,3}$的前$j-1$个元素和$c_{j-1,1}$中的元素是相同的，而最后的元素分别是1和0. 由于式(530)中的相加项数是4的倍数，而且每4个一组可以像前4项一样处理，因此证明了结论.

练习[240]答案 我们有

$$
\begin{aligned}
\langle \boldsymbol{X}^{(1)},\boldsymbol{R}^{(1)}\rangle &=(\boldsymbol{X}^{(1)})^{\mathrm{T}}\boldsymbol{R}^{(1)}\\
&=\langle \boldsymbol{X},\boldsymbol{d}_{\gamma_0}\rangle\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}(\boldsymbol{X}-\boldsymbol{X}^{(1)})\\
&=\langle \boldsymbol{X},\boldsymbol{d}_{\gamma_0}\rangle(\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}\boldsymbol{X}-\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}\boldsymbol{X}^{(1)}),
\end{aligned}
$$

所以，如果$\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}\boldsymbol{X}^{(1)}$时结论成立. 为了证明这是真的，注意

$$
\begin{aligned}
\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}\boldsymbol{X}^{(1)}&=\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}(\langle \boldsymbol{X},\boldsymbol{d}_{\gamma_0}\rangle\boldsymbol{d}_{\gamma_0})\\
&=\langle \boldsymbol{X},\boldsymbol{d}_{\gamma_0}\rangle\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}\boldsymbol{d}_{\gamma_0}\\
&=\langle \boldsymbol{X},\boldsymbol{d}_{\gamma_0}\rangle=\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}\boldsymbol{X},
\end{aligned}
$$

因为$\boldsymbol{d}_{\gamma_0}^{\mathrm{T}}\boldsymbol{d}_{\gamma_0}=\|\boldsymbol{d}_{\gamma_0}\|^2=1$. 我们有

$$
\begin{aligned}
\|\boldsymbol{X}\|^2&=\|\boldsymbol{X}^{(1)}+\boldsymbol{R}^{(1)}\|^2\\
&=\|\boldsymbol{X}^{(1)}\|^2+\|\boldsymbol{R}^{(1)}\|^2+2\langle \boldsymbol{X}^{(1)},\boldsymbol{R}^{(1)}\rangle\\
&=\|\boldsymbol{X}^{(1)}\|^2+\|\boldsymbol{R}^{(1)}\|^2;
\end{aligned}
$$

另外

$$
\begin{aligned}
\|\boldsymbol{X}^{(1)}\|^2&=\|\langle \boldsymbol{X},\boldsymbol{d}_{\gamma_0}\rangle\boldsymbol{d}_{\gamma_0}\|^2\\
&=|\langle \boldsymbol{X},\boldsymbol{d}_{\gamma_0}\rangle|^2\|\boldsymbol{d}_{\gamma_0}\|^2\\
&=|\langle \boldsymbol{X},\boldsymbol{d}_{\gamma_0}\rangle|^2,
\end{aligned}
$$

这就完成了练习.

练习[261]答案 我们有

$$
E\{(X-a)^2\}=E\{X^2\}-2aE\{X\}+a^2\equiv q(a),
$$

这是个关于a的二次函数. 为了确定这个函数的极值，我们对于上式关于a求导数，得到并且置结果为0有

$$
\frac{\mathrm{d}q(a)}{\mathrm{d}a}=-2E\{X\}+2a=0
$$

(由于二阶导数是2，这个极值事实上是最小值). 上式等价于$a=E\{X\}$.

练习[262a]答案 令 $u_0(x_0, x_1)\equiv x_0$ 和 $u_1(x_0, x_1)\equiv x_0+x_1$．这个变换的雅可比(Jacobi)行列式是

$$J=\begin{vmatrix}1 & 0\\ 1 & 1\end{vmatrix}=1.$$

而有 $y_0=u_0(x_0, x_1)$ 和 $y_1=u_1(x_0, x_1)$，我们可以得到它的逆函数 $v_0(y_0, y_1)=y_0$ 和 $v_1(y_0, y_1)=y_1-y_0$．因此，我们有

$$f_{Y_0,Y_1}(y_0,y_1)=f_{X_0,X_1}(v_0(y_0,y_1),v_1(y_0,y_1))=f_{X_0,X_1}(y_0,y_1-y_0).$$

练习[262b]答案 回忆方阵的迹是它的对角元素的和，因此有

$$\operatorname{tr}\{\Sigma_X\}=\sum_{t=0}^{N-1}\operatorname{var}\{X_t\}.$$

因此我们需要证明 $\operatorname{tr}\{\Sigma_Y\}=\operatorname{tr}\{\Sigma_X\}$．如果 A 和 B 是使乘积 AB 是方阵的矩阵，在矩阵理论中一个标准的结果就是 $\operatorname{tr}\{AB\}=\operatorname{tr}\{BA\}$(例如见 Graybill，1983，9.1 节)．于是，由于$\mathcal{O}^T\mathcal{O}=I_N$，

$$\operatorname{tr}\{\Sigma_Y\}=\operatorname{tr}\{\mathcal{O}\Sigma_X\mathcal{O}^T\}=\operatorname{tr}\{\Sigma_X\mathcal{O}^T\mathcal{O}\}=\operatorname{tr}\{\Sigma_X I_N\}=\operatorname{tr}\{\Sigma_X\}$$

对于正交变换成立．

练习[263]答案 如果 $\mu_X=\mathbf{0}$，那么由式(262b)得出 $\mu_Y=\mathcal{O}\mu_X=\mathbf{0}$；同样，如果 $\Sigma_X=\sigma^2 I_N$，那么由式(262c)得出，由于$\mathcal{O}\mathcal{O}^T=I_N$ 规范正交，$\Sigma_Y=\mathcal{O}\Sigma_X\mathcal{O}^T=\sigma^2\mathcal{O}I_N\mathcal{O}^T=\sigma^2\mathcal{O}\mathcal{O}^T=\sigma^2 I_N$．由于 $\mathbf{X}$ 的多元高斯分布也隐含着 $\mathbf{Y}$ 的多元高斯分布，我们能够推出，因此 $\mathbf{X}$ 和 $\mathbf{Y}$ 的统计性质是等同的．最后，如果 $\mu_X\neq\mathbf{0}$，我们仍有 $\Sigma_Y=\Sigma_X$，但是现在 μ_Y 不需要等于 μ_X．例如，假设 $\mu_X=[\mu, \mu]^T$，考虑它的哈尔离散傅里叶矩阵

$$W\equiv\begin{bmatrix}-\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix},$$

对这个矩阵，

$$\mu_Y=W\mu_X=\begin{bmatrix}0\\ \mu\sqrt{2}\end{bmatrix}.$$

练习[265]答案 使用 $r_d(\theta)$的定义，我们有

$$\begin{aligned}E_\theta\{r_d(\theta)\}&=\int r_d(\theta_0)f_\theta(\theta_0)\mathrm{d}\theta_0\\&=\int\left(\int\ell(\mathrm{d}(x),\theta_0)f_{X|\theta=\theta_0}(x)\mathrm{d}x\right)f_\theta(\theta_0)\mathrm{d}\theta_0\\&=\iint\ell(\mathrm{d}(x),\theta_0)f_{X|\theta=\theta_0}(x)f_\theta(\theta_0)\mathrm{d}x\mathrm{d}\theta_0.\end{aligned}$$

由式(260)，我们有 $f_{X|\theta=\theta_0}(x)f_\theta(\theta_0)=f_{X,\theta}(x, \theta_0)$；而 $f_{X,\theta}(x, \theta_0)=f_{\theta|X=x}(\theta_0)f_X(x)$，所以有

$$\begin{aligned}E_\theta\{r_d(\theta)\}&=\iint\ell(\mathrm{d}(x),\theta_0)f_{\theta|X=x}(\theta_0)f_X(x)\mathrm{d}x\mathrm{d}\theta_0\\&=\int\left(\int\ell(\mathrm{d}(x),\theta_0)f_{\theta|X=x}(\theta_0)\mathrm{d}\theta_0\right)f_X(x)\mathrm{d}x.\end{aligned}$$

练习[288a]答案 首先注意到，因为 $X_0\equiv 0$ 和 $E\{\varepsilon_u\}=0$ 对于所有的 u 成立，我们有 $E\{X_0\}=0$，而对于 $t>0$ 和 $t<0$ 我们分别有

$$E\{X_t\}=\sum_{u=1}^{t}E\{\varepsilon_u\}=0 \quad \text{和} \quad E\{X_t\}=-\sum_{u=0}^{|t|-1}E\{\varepsilon_{-u}\}=0$$

对于所有 $|t|>0$ 成立．因此随机游动过程的第一个矩不依赖于 t(这时要求是平稳过程)．现在让我们考虑 $\{X_t\}$ 的方差，我们有 $\mathrm{var}\{X_0\}=0$，并且对于 $t>0$，

$$\begin{aligned}\mathrm{var}\{X_t\}&=E\{X_t^2\}\\&=E\Big\{\Big(\sum_{u=1}^{t}\varepsilon_u\Big)\Big(\sum_{u'=1}^{t}\varepsilon_{u'}\Big)\Big\}\\&=\sum_{u=1}^{t}\sum_{u'=1}^{t}E\{\varepsilon_u\varepsilon_{u'}\}=t\sigma_\varepsilon^2,\end{aligned}$$

由于 $E\{\varepsilon_u\varepsilon_{u'}\}$ 非零仅当 $u=u'$ 成立，这恰是在二重和中希望的 $|t|$．由于 $\mathrm{var}\{X_t\}$ 依赖于 t，随机游动过程 $\{X_t\}$ 是非平稳的．为了确定(b)部分，设 $Y_t\equiv X_t-X_{t-1}$ 为 $\{X_t\}$ 的一阶向后差分．对 $t=1$，我们有 $Y_1=X_1-X_0=\varepsilon_1$；对 $t>1$，我们有

$$Y_t=\sum_{u=1}^{t}\varepsilon_u-\sum_{u=1}^{t-1}\varepsilon_u=\varepsilon_t;$$

对 $t=0$，我们有 $Y_0=X_0-X_{-1}=\varepsilon_0$；并且 $t<0$，我们有

$$Y_t=-\sum_{u=0}^{|t|-1}\varepsilon_{-u}+\sum_{u=0}^{|t-1|-1}\varepsilon_{-u}=\varepsilon_t.$$

于是 $\{Y_t\}$ 是白噪声过程 $\{\varepsilon_t\}$，所以 $\{X_t\}$ 的一阶向后差分是平稳过程．为了确定(c)部分，注意到，由于白噪声过程 $\{Y_t\}$ 的谱密度函数由 $S_Y(f)=\sigma_\varepsilon^2$ 给出，由此得出 $\{X_t\}$ 的谱密度函数由

$$\begin{aligned}S_X(f)&=\frac{S_Y(f)}{4\sin^2(\pi f)}\\&=\frac{\sigma_\varepsilon^2}{4\sin^2(\pi f)}\approx\frac{\sigma_\varepsilon^2}{4\pi^2f^2}\end{aligned}$$

给出，其中对小的 f 近似成立(因为 $\sin(x)\approx x$ 对于小的 x 成立)．

练习[288b]答案 由于对于所有 t $E\{U_t\}=0$，由此得出

$$\begin{aligned}E\{X_t\}&=E\{a_0+a_1t+U_t\}\\&=a_0+a_1t,\end{aligned}$$

这依赖于 t，所以 $\{X_t\}$ 是非平稳过程．为了建立(b)，设 $\{s_{U,\tau}\}$ 表示对 $\{U_t\}$ 的自协方差序列．注意一阶向后差分

$$\begin{aligned}Y_t&\equiv X_t-X_{t-1}\\&=a_0+a_1t+U_t-(a_0+a_1(t-1)+U_{t-1})\\&=a_1+U_t-U_{t-1}\end{aligned}$$

具有均值

$$\begin{aligned}E\{Y_t\}&=E\{a_1+U_t-U_{t-1}\}\\&=a_1+E\{U_t\}-E\{U_{t-1}\}=a_1\neq 0\end{aligned}$$

以及随机变量和 $Y_{t+\tau}$ 之间的协方差

$$\begin{aligned}\operatorname{cov}\{Y_t, Y_{t+\tau}\} &= E\{(Y_t - a_1)(Y_{t+\tau} - a_1)\} \\ &= E\{(U_t - U_{t-1})(U_{t+\tau} - U_{t+\tau-1})\} \\ &= E\{U_t U_{t+\tau}\} - E\{U_t U_{t+\tau-1}\} - E\{U_{t-1} U_{t+\tau}\} + E\{U_{t-1} U_{t+\tau-1}\} \\ &= s_{U,\tau} - s_{U,\tau-1} - s_{U,\tau+1} + s_{U,\tau} \\ &= 2s_{U,\tau} - s_{U,\tau-1} - s_{U,\tau+1} \equiv s_{Y,\tau}.\end{aligned}$$

由于 $E\{Y_t\}$ 和 $\operatorname{cov}\{Y_t, Y_{t+\tau}\}$ 都是独立于 t 并且是有限的，过程 $E\{Y_t\}$ 是平稳的并具有非零均值．为了建立(c)，设 $\{Z_t\}$ 为 $\{Y_t\}$ 一阶向后差分，这和 $\{X_t\}$ 二阶向后差分是一样的．我们有

$$\begin{aligned}E\{Z_t\} &= E\{Y_t - Y_{t-1}\} \\ &= E\{Y_t\} - E\{Y_{t-1}\} = 0,\end{aligned}$$

我们能使用如前论证得到

$$\begin{aligned}\operatorname{cov}\{Z_t, Z_{t+\tau}\} &= 2s_{Y,\tau} - s_{Y,\tau-1} - s_{Y,\tau+1} \\ &= 2(2s_{U,\tau} - s_{U,\tau-1} - s_{U,\tau+1}) \\ &\quad - (2s_{U,\tau-1} - s_{U,\tau-2} - s_{U,\tau}) \\ &\quad - (2s_{U,\tau+1} - s_{U,\tau} - s_{U,\tau+2}) \\ &= 6s_{U,\tau} - 4s_{U,\tau-1} - 4s_{U,\tau+1} + s_{U,\tau-2} + s_{U,\tau+2},\end{aligned}$$

上式依赖于 τ，而不依赖 t，因此 $\{Z_t\}$ 是具有零均值的平稳过程．

练习[289]答案 从式(299a)中我们有

$$\begin{aligned}E\{\hat{\sigma}_Y^2\} &= \frac{1}{N}\sum_{t=0}^{N-1} E\{[(Y_t - \mu_Y) - (\bar{Y} - \mu_Y)]^2\} \\ &= \frac{1}{N}\sum_{t=0}^{N-1} E\{(Y_t - \mu_Y)^2 - 2(Y_t - \mu_Y)(\bar{Y} - \mu_Y) + (\bar{Y} - \mu_Y)^2\} \\ &= \sigma_Y^2 + \operatorname{var}\{\bar{Y}\} - \frac{2}{N}E\left\{(\bar{Y} - \mu_Y)\sum_{t=0}^{N-1}(Y_t - \mu_Y)\right\} \\ &= \sigma_Y^2 + \operatorname{var}\{\bar{Y}\} - \frac{2}{N}E\{(\bar{Y} - \mu_Y)(N\bar{Y} - N_{\mu_Y})\} \\ &= \sigma_Y^2 - \operatorname{var}\{\bar{Y}\}.\end{aligned}$$

练习[300]答案 设 $\bar{Y}_N \equiv \frac{1}{N}\sum_{t=0}^{N-1} Y_t$．用这个记号，我们需要证明 $\operatorname{var}\{\bar{Y}_N\} = \sigma_Y^2 N^{2H-2}$．由于 $\operatorname{var}\{\bar{Y}_1\} = \operatorname{var}\{Y_0\} = \sigma_Y^2$，结果在 $N=1$ 时成立．假设它在 $N-1$ 时也成立；即 $\operatorname{var}\{\bar{Y}_{N-1}\} = \sigma_Y^2 (N-1)^{2H-2}$．反复地使用练习[7.2]的(d)部分，则有

$$\begin{aligned}\operatorname{var}\{\bar{Y}_N\} &= \operatorname{var}\left\{\frac{N-1}{N}\bar{Y}_{N-1} + \frac{1}{N}Y_{N-1}\right\} \\ &= \operatorname{var}\left\{\frac{N-1}{N}\bar{Y}_{N-1}\right\} + \operatorname{var}\left\{\frac{1}{N}Y_{N-1}\right\} + 2\operatorname{cov}\left\{\frac{N-1}{N}\bar{Y}_{N-1}, \frac{1}{N}Y_{N-1}\right\} \\ &= \frac{(N-1)^2}{N^2}\sigma_Y^2 (N-1)^{2H-2} + \frac{\sigma_Y^2}{N^2} + \frac{2}{N^2}\operatorname{cov}\left\{\sum_{t=0}^{N-2} Y_t, Y_{N-1}\right\}\end{aligned}$$

$$= \frac{1}{N^2}\Big[\sigma_Y^2(N-1)^{2H} + \sigma_Y^2 + 2\sum_{t=0}^{N-2}\operatorname{cov}\{Y_t, Y_{N-1}\}\Big]$$

$$= \frac{1}{N^2}\Big[\sigma_Y^2(N-1)^{2H} + \sigma_Y^2 + 2\sum_{\tau=1}^{N-1}s_{Y,\tau}\Big]$$

$$= \frac{\sigma_Y^2}{N^2}\Big[(N-1)^{2H} + 1 + \sum_{\tau=1}^{N-1}(|\tau+1|^{2H} - 2|\tau|^{2H} + |\tau-1|^{2H})\Big]$$

$$= \frac{\sigma_Y^2}{N^2}\Big[(N-1)^{2H} + 1 + \sum_{\tau=2}^{N}\tau^{2H} - 2\sum_{\tau=1}^{N-1}\tau^{2H} + \sum_{\tau=1}^{N-2}\tau^{2H}\Big]$$

$$= \frac{\sigma_Y^2}{N^2}[(N-1)^{2H} + 1 + N^{2H} - (N-1)^{2H} - 1] = \sigma_Y^2 N^{2H-2},$$

这就完成了证明.

练习[302]答案 假设式(302)对一些 $J_0 \geqslant 2$ 成立：

$$\operatorname{var}\{Y_t\} = \operatorname{var}\{\bar{V}_{J_0-1,t}\} + \sum_{j=1}^{J_0-1}\nu_Y^2(\tau_j).$$

现在

$$\operatorname{var}\{\bar{V}_{J_0-1,t}\} = \int_{-1/2}^{1/2}\widetilde{\mathcal{G}}_{J_0-1}(f)S_Y(f)\mathrm{d}f$$

$$= \int_{-1/2}^{1/2}\Big(\prod_{l=0}^{J_0-2}\widetilde{\mathcal{G}}(2^l f)\Big)S_Y(f)\mathrm{d}f$$

$$= \int_{-1/2}^{1/2}\Big([\widetilde{\mathcal{G}}(2^{J_0-1}f) + \widetilde{\mathcal{H}}(2^{J_0-1}f)]\prod_{l=0}^{J_0-2}\widetilde{\mathcal{G}}(2^l f)\Big)S_Y(f)\mathrm{d}f$$

$$= \int_{-1/2}^{1/2}[\widetilde{\mathcal{G}}_{J_0}(f) + \widetilde{\mathcal{H}}_{J_0}(f)]S_Y(f)\mathrm{d}f$$

$$= \operatorname{var}\{\bar{V}_{J_0,t}\} + \nu_Y^2(\tau_{J_0}),$$

所以，我们有

$$\operatorname{var}\{Y_t\} = \operatorname{var}\{\bar{V}_{J_0,t}\} + \nu_Y^2(\tau_{J_0}) + \sum_{j=1}^{J_0-1}\nu_Y^2(\tau_j)$$

$$= \operatorname{var}\{\bar{V}_{J_0,t}\} + \sum_{j=1}^{J_0}\nu_Y^2(\tau_j).$$

练习[304]答案 对于 $j=1$ 情况的证明可以很容易地延伸到一般的 j，如果我们可以建立证明，滤波器 $\{\tilde{h}_{j,l}\}$ 是两部分级联滤波器的等价滤波器：第一部分是 d 阶的向后差分滤波器(这个将 $\{X_t\}$ 变为平稳过程 $\{Y_t\}$)，第二部分为具有有限宽度的滤波器. 很容易从式(169b)和对 $j=1$ 的证明得到，$\{\tilde{h}_{j,l}\}$，$j\geqslant 2$ 的平方增益函数 $\widetilde{\mathcal{H}}_j^{(D)}(\cdot)$ 的表达式为

$$\widetilde{\mathcal{H}}_j^{(D)}(f) = \widetilde{\mathcal{H}}^{(D)}(2^{j-1}f)\,\widetilde{\mathcal{G}}_{j-1}^{(D)}(f)$$

$$= \mathcal{D}^{\frac{L}{2}}(2^{j-1}f)\,\widetilde{\mathcal{A}}_L(2^{j-1}f)\,\widetilde{\mathcal{G}}_{j-1}^{(D)}(f).$$

应用$\mathcal{D}(\cdot)$的定义和指示中三角恒等式，对于$j\geqslant 2$

$$\begin{aligned}\mathcal{D}(2^{j-1}f) &= 4\sin^2(\pi 2^{j-1}f)\\ &= 4\sin^2(\pi 2^{j-2}f)[4\cos^2(\pi 2^{j-2}f)]\\ &= 4\sin^2(\pi 2^{j-3}f)[4\cos^2(\pi 2^{j-3}f)][4\cos^2(\pi 2^{j-2}f)]\\ &\vdots\\ &= \mathcal{D}(f)\prod_{k=0}^{j-2}4\cos^2(\pi 2^k f),\end{aligned}$$

其中 $4\cos^2(\pi f)$定义向后和滤波器$\{1,1\}$的平方增益函数．因此，由$\mathcal{D}^{\frac{L}{2}}(2^{j-1}f)$定义的平方增益函数可以分成两个滤波器，第一个是 $L/2$ 阶向后差分滤波器．由于 $L/2\geqslant d$，我们现在能写

$$\widetilde{\mathcal{H}}_j^{(\mathrm{D})}(f) = \mathcal{D}^d(f)\,\mathcal{B}_j(f),$$

其中

$$\mathcal{B}_j(f) \equiv \mathcal{D}^{\frac{L}{2}-d}(f)\Big[\prod_{k=0}^{j-2}4\cos^2(\pi 2^k f)\Big]^{\frac{L}{2}}\,\widetilde{\mathcal{A}}_L(2^{j-1}f)\,\widetilde{\mathcal{G}}_{j-1}^{(\mathrm{D})}(f) \tag{536}$$

必须是一个有限宽度滤波器的平方增益函数，因为级联滤波器的等价滤波器$\{\tilde{h}_{j,l}\}$具有有限的宽度．

练习[305]答案 用练习[304]建立的结果说明，我们能写

$$\overline{W}_{j,t} = \sum_{l=0}^{L_j-d-1} b_{j,l}Y_{t-l},$$

其中$\{b_{j,l}\}$滤波器它的由式(536)定义的平方增益函数$\mathcal{B}_j(\cdot)$，$\{Y_t\}$是具有均值 μ_Y 的平稳过程，它是由 d 阶向后差分滤波器滤波$\{X_t\}$形成的．设 $B_j(\cdot)$表示$\{b_{j,l}\}$的传递函数(因此 $|B_j(f)|^2=\mathcal{B}_j(f)$)，则有

$$\begin{aligned}E\{\overline{W}_{j,t}\} &= \sum_{l=0}^{L_j-d-1} b_{j,l}E\{Y_{t-l}\}\\ &= \mu_Y\sum_{l=0}^{L_j-d-1} b_{j,l} = \mu_Y B_j(0),\end{aligned}$$

所以当 $\mu_Y=0$ 时 $E\{\overline{W}_{j,t}\}=0$．由于$\mathcal{D}(0)=4\sin^2(\pi\cdot 0)=0$，式(536)告诉我们$\mathcal{B}_j(0)=0$ 并且当 $L/2>d$ 时 $B_j(0)=0$，在这种情况下，不管 μ_Y 是多少，$E\{\overline{W}_{j,t}\}=0$．余下证明当 $\mu_Y\neq 0$ 和 $L/2=d$ 时 $E\{\overline{W}_{j,t}\}\neq 0$，如果我们可以证明 $B_j(0)\neq 0$，或者等价地，$\mathcal{B}_j(0)>0$ 这个结论成立．当 $L/2=d$ 时

$$\begin{aligned}\mathcal{B}_j(0) &= \widetilde{\mathcal{A}}_L(0)\,\widetilde{\mathcal{G}}_{j-1}^{(\mathrm{D})}(0) = \widetilde{\mathcal{A}}_L(0)\\ &= \frac{1}{2^L}\sum_{l=0}^{\frac{L}{2}-1}\binom{\frac{L}{2}-1+l}{l} = \begin{cases}128/512, & L=2;\\ 96/512, & L=4;\\ 80/512, & L=6;\text{等等},\end{cases}\end{aligned}$$

其中我们用到事实

$$\widetilde{\mathcal{G}}_{j-1}^{(\mathrm{D})}(0) = \prod_{l=0}^{j-2}\widetilde{\mathcal{G}}(0) = 1.$$

这就建立了练习的结论．

练习[312]答案 假设谱密度函数估计量$\hat{S}_j(\cdot)$为

$$\frac{\eta \hat{S}_j(f)}{S_j(f)} \stackrel{\mathrm{d}}{=} \chi^2_\eta, \quad 0 < |f| < 1/2.$$

由于$E\{\chi^2_\eta\}=\eta$且$\mathrm{var}\{\chi^2_\eta\}=2\eta$，有

$$E\left\{\frac{\eta \hat{S}_j(f)}{S_j(f)}\right\} = \eta, \quad \text{即} \quad E\{\hat{S}_j(f)\} = S_j(f),$$

且

$$\mathrm{var}\left\{\frac{\eta \hat{S}_j(f)}{S_j(f)}\right\} = 2\eta, \quad \text{即} \quad \mathrm{var}\{\hat{S}_j(f)\} = \frac{2S_j^2(f)}{\eta}.$$

对于任一有着均值$E\{U\}$的随机变量U，有

$$\mathrm{var}\{U\} \equiv E\{(U - E\{U\})^2\} = E\{U^2\} - (E\{U\})^2,$$

所以$E\{U^2\} = \mathrm{var}\{U\} + (E\{U\})^2$. 因此

$$\begin{aligned} E\{\hat{S}_j^2(f)\} &= \mathrm{var}\{\hat{S}_j(f)\} + (E\{\hat{S}_j(f)\})^2 \\ &= \frac{2S_j^2(f)}{\eta} + S_j^2(f) \\ &= S_j^2(f)\left(\frac{2}{\eta} + 1\right). \end{aligned} \tag{537}$$

对于周期图，有$\eta=2$，所以式(537)告诉我们

$$E\{[\hat{S}_j^{(\mathrm{p})}(f)]^2\} = 2S_j^2(f).$$

因此我们近似有

$$\begin{aligned} E\{\hat{A}_j\} &= \frac{1}{2}\int_{-1/2}^{1/2} E\{[\hat{S}_j^{(\mathrm{p})}(f)]^2\}\mathrm{d}f \\ &= \frac{1}{2}\int_{-1/2}^{1/2} 2S_j^2(f)\mathrm{d}f \\ &= \int_{-1/2}^{1/2} S_j^2(f)\mathrm{d}f = A_j \end{aligned}$$

(频率$f=0$，$\pm 1/2$的值不影响积分)；即$\hat{A}_j$是A_j的一个近似无偏估计量.

练习[313a]答案 由于

$$E\{\hat{\nu}_X^2(\tau_j)\} = E\{a\chi^2_{\eta_1}\} = a\eta_1$$

和

$$\mathrm{var}\{\hat{\nu}_X^2(\tau_j)\} = \mathrm{var}\{a\chi^2_{\eta_1}\} = 2a^2\eta_1,$$

有

$$\frac{2(E\{\hat{\nu}_X^2(\tau_j)\})^2}{\mathrm{var}\{\hat{\nu}_X^2(\tau_j)\}} = \frac{2a^2\eta_1^2}{2a^2\eta_1} = \eta_1.$$

利用$E\{\hat{\nu}_X^2(\tau_j)\}$和$\mathrm{var}\{\hat{\nu}_X^2(\tau_j)\}$的大样本近似产生

$$\eta_1 = \frac{2\nu_X^4(\tau_j)}{2A_j/M_j} = \frac{M_j\nu_X^4(\tau_j)}{A_j}.$$

作为补充，也注意

$$a = \frac{E\{\hat{\nu}_X^2(\tau_j)\}}{\eta_1} = \frac{A_j}{M_j\nu_X^2(\tau_j)}.$$

练习[313b]答案 由 $Q_\eta(p)$的定义，(对于 $0\leqslant p\leqslant 1/2$)我们有

$$P\left[Q_\eta(p)\leqslant \chi_\eta^2\leqslant Q_\eta(1-p)\right]=1-2p.$$

由于通过假设 $\eta\hat{\nu}_X^2(\tau_j)/\nu_X^2(\tau_j)$有一个 χ_η^2 随机变量分布，有

$$P\left[Q_\eta(p)\leqslant \frac{\eta\hat{\nu}_X^2(\tau_j)}{\nu_X^2(\tau_j)}\leqslant Q_\eta(1-p)\right]=1-2p.$$

方括号中的事件发生当且仅当事件

$$Q_\eta(p)\leqslant \frac{\eta\hat{\nu}_X^2(\tau_j)}{\nu_X^2(\tau_j)} \quad \text{和} \quad \frac{\eta\hat{\nu}_X^2(\tau_j)}{\nu_X^2(\tau_j)}\leqslant Q_\eta(1-p)$$

同时发生；这两个事件等价于

$$\nu_X^2(\tau_j)\leqslant \frac{\eta\hat{\nu}_X^2(\tau_j)}{Q_\eta(p)} \quad \text{和} \quad \frac{\eta\hat{\nu}_X^2(\tau_j)}{Q_\eta(1-p)}\leqslant \nu_X^2(\tau_j).$$

综合这两个事件等价于单个事件

$$\frac{\eta\hat{\nu}_X^2(\tau_j)}{Q_\eta(1-p)}\leqslant \nu_X^2(\tau_j)\leqslant \frac{\eta\hat{\nu}_X^2(\tau_j)}{Q_\eta(p)},$$

这样就给出了练习中规定的置信区间.

练习[314a]答案 对$\hat{S}_j^{(p)}(f_k)$分布的大样本近似告诉我们 $E\{\hat{S}_j^{(p)}(f_k)\}=S_j(f_k)=aC_j(f_k)$和 $\operatorname{var}\{\hat{S}_j^{(p)}(f_k)\}=S_j^2(f_k)=a^2C_j^2(f_k)$. 使用式(314a)，有

$$\begin{aligned}E\{\hat{\nu}_X^2(\tau_j)\}&\approx \frac{2}{M_j}\sum_{k=1}^{\lfloor (M_j-1)/2\rfloor}E\{\hat{S}_j^{(p)}(f_k)\}\\&=\frac{2a}{M_j}\sum_{k=1}^{\lfloor (M_j-1)/2\rfloor}C_j(f_k);\end{aligned}$$

再利用这个公式，以及相互独立随机变量的和的方差等于随机变量方差的和的事实，我们有

$$\operatorname{var}\{\hat{\nu}_X^2(\tau_j)\}\approx \frac{4a^2}{M_j^2}\sum_{k=1}^{\lfloor (M_j-1)/2\rfloor}C_j^2(f_k).$$

使用练习[313a]的结果得到

$$\begin{aligned}\eta_2&=\frac{2(E\{\hat{\nu}_X^2(\tau_j)\})^2}{\operatorname{var}\{\hat{\nu}_X^2(\tau_j)\}}\\&=\frac{2\left(\frac{2a}{M_j}\sum_{k=1}^{\lfloor (M_j-1)/2\rfloor}C_j(f_k)\right)^2}{\frac{4a^2}{M_j^2}\sum_{k=1}^{\lfloor (M_j-1)/2\rfloor}C_j^2(f_k)},\end{aligned}$$

从它可以得到要求的结果.

练习[314b]答案 用假设，我们能写 $S_j(f)\approx aC_j(f)$，其中

$$C_j(f)\equiv\begin{cases}1, & \frac{1}{2^{j+1}}<|f|\leqslant\frac{1}{2^j};\\ 0, & \text{其他}.\end{cases}$$

现在，$\hat{\nu}_X^2(\tau_j)$是有着均值 $\nu_X^2(\tau_j)$和大样本方差 $2A_j/M_j$ 的渐近正态. 其中由式(307a)，

$$\begin{aligned}A_j&=\int_{-1/2}^{1/2}S_j^2(f)\,\mathrm{d}f\\&\approx 2\int_{-1/2}^{1/2}a^2C_j^2(f)\,\mathrm{d}f\\&=2\int_{1/2^{j+1}}^{1/2^j}a^2\,\mathrm{d}f=\frac{a^2}{2^j}.\end{aligned}$$

因此可以得 $\mathrm{var}\{\hat{\nu}_X^2(\tau_j)\}\approx a^2/(2^{j-1}M_j)$. 由式(305)，有

$$\begin{aligned}\nu_X^2(\tau_j)&=\int_{-1/2}^{1/2}S_j(f)\,\mathrm{d}f\\&\approx\int_{-1/2}^{1/2}aC_j(f)\,\mathrm{d}f\\&=2\int_{1/2^{j+1}}^{1/2^j}a\,\mathrm{d}f=\frac{a}{2^j},\end{aligned}$$

并且因此 $E\{\hat{\nu}_X^2(\tau_j)\}\approx a/2^j$. 从练习[313a]，

$$\eta=\frac{2(E\{\hat{\nu}_X^2(\tau_j)\})^2}{\mathrm{var}\{\hat{\nu}_X^2(\tau_j)\}}\approx\frac{2(a/2^j)^2}{a^2/(2^{j-1}M_j)}=\frac{M_j}{2^j}.$$

如果 $M_j/2^j<1$，可以确定一个较好的近似是 1，因为用于形成$\hat{\nu}_X^2(\tau_j)$的和总由至少一个有着零均值的正态分布随机变量组. 这就导致了练习要求建立的 $\eta_3=\max\{M_j/2^j,1\}$.

练习[315]答案 定义 $\eta\equiv 2N_B$ 和

$$\widetilde{A}_j\equiv C\int_{-1/2}^{1/2}[\hat{S}_{(\mathrm{ss}),j}(f)]^2\,\mathrm{d}f,$$

其中 C 有待确定. 由 $\eta\hat{S}_{(\mathrm{ss}),j}(f)/S_j(f)$分布的假设，我们有 $E\{\hat{S}_{(\mathrm{ss}),j}(f)\}=S_j(f)$和 $\mathrm{var}\{\hat{S}_{(\mathrm{ss}),j}(f)\}=2S_j^2(f)/\eta$. 由于对任一的随机变量 $E\{U^2\}=\mathrm{var}\{U\}+(E\{U\})^2$，我们得到

$$\begin{aligned}E\{\widetilde{A}_j\}&=C\int_{-1/2}^{1/2}E\{[\hat{S}_{(\mathrm{ss}),j}(f)]^2\}\,\mathrm{d}f\\&=C\int_{-1/2}^{1/2}S_j^2(f)\left(\frac{2}{\eta}+1\right)\mathrm{d}f\\&=C\left(\frac{2}{\eta}+1\right)A_j,\end{aligned}$$

所以我们通过取

$$C\equiv\frac{1}{\frac{2}{\eta}+1}=\frac{\eta}{2+\eta}$$

得到一个无偏估计量. 随着 N_B(并且因此 η)增长至无穷，我们有 $C\to 1$，所以估计量变成

$$\widetilde{A}_j=\int_{-1/2}^{1/2}[\hat{S}_{(\mathrm{ss}),j}(f)]^2\,\mathrm{d}f,$$

从 A_j 的定义可知它是合理的.

练习[320a]答案 为了对于尺度 τ_j 获得一个无偏极大重叠离散小波变换小波方差估计量，我们必须有 $M_j=N-L_j+1\geqslant 1$，其中 $L_j=(2^j-1)(L-1)+1$. 对于 $N=1026$，当 $L=2, 4, 6$ 时，不等式成立的最大 j 分别为 $j=10, 8, 7$.

练习[320b]答案 j 层极大重叠小波变换哈尔小波滤波器有一个脉冲响应序列 $\{\tilde{h}_l\}$ 包含以下 $L_j=2^j$ 个非零值：

$$\underbrace{\frac{1}{2^j},\cdots,\frac{1}{2^j}}_{2^{j-1}\text{个}},\underbrace{-\frac{1}{2^j},\cdots,-\frac{1}{2^j}}_{2^{j-1}\text{个}}.$$

因此对 $\{X_t\}$，相应的极大重叠小波变换系数由

$$\begin{aligned}\widetilde{W}_{j,t}&=\sum_{l=0}^{L_j-1}\tilde{h}_l X_{t-l}\\&=\frac{1}{2^j}\sum_{l=0}^{2^{j-1}-1}X_{t-l}-X_{t-2^{j-1}-l}\\&=\frac{1}{2^j}\sum_{l=0}^{2^{j-1}-1}\left[(X'_{t-l}+a_0+a_1(t-l))-(X'_{t-2^{j-1}-l}+a_0+a_1(t-2^{j-1}-l))\right]\\&=\frac{1}{2^j}\sum_{l=0}^{2^{j-1}-1}(X'_{t-l}-X'_{t-2^{j-1}-l}+a_1 2^{j-1})\\&=\widetilde{W}'_{j,t}+\frac{a_1}{2}\tau_j\end{aligned}$$

给出. 所以

$$\begin{aligned}E\{\hat{\nu}_X^2(\tau_j)\}&=E\{\widetilde{W}_{j,t}^2\}\\&=E\{[\widetilde{W}'_{j,t}]^2\}+a_1\tau_jE\{\widetilde{W}'_{j,t}\}+\frac{a_1^2}{4}\tau_j^2\\&=\nu_{X'}^2(\tau_j)+\frac{a_1^2}{4}\tau_j^2\end{aligned}$$

(练习[305]第一个结果告诉我们 $E\{\widetilde{W}'_{j,t}\}=0$). 如果上面的第二项支配第一项，那么有

$$\nu_X(\tau_j)\approx\frac{|a_1|}{2}\tau_j,\quad \text{所以}\ \log_{10}(\nu_X(\tau_j))\approx-\log_{10}(2)+\log_{10}(|a_1|)+\log_{10}(\tau_j),$$

因此作为 $\log_{10}(\tau_j)$ 为函数的 $\log_{10}(\nu_X(\tau_j))$ 的图近似为有单位斜率的直线. 我们应该也有 $\nu_X(\tau_1)\approx|a_1|/2$. 从图 318 上部得到关于 a_1 的一个粗略估计是 $-150\,000/1000=-150$ 毫微秒，所以我们应该有 $\nu_X(\tau_1)\doteq 75$ 毫微秒；事实上 $\hat{\nu}_X(\tau_1)\doteq 71.38$ 毫微秒，所以图 319 哈尔小波方差的现象几乎可以完全解释为根据 $\{X_t\}$ 的近似线性漂移.

练习[322]答案 就像对练习[320b] 的解，我们可以将 $\{\overline{Y}_t(\tau_1)\}$ 的 j 层极大重叠小波变换哈尔小波系数过程表示为

$$\begin{aligned}\widetilde{W}_{j,t}&=\frac{1}{2^j}\sum_{l=0}^{2^{j-1}-1}\overline{Y}_{t-l}(\tau_1)-\overline{Y}_{t-2^{j-1}-l}(\tau_1)\\&=\frac{1}{2\tau_j}\sum_{l=0}^{\tau_j-1}\overline{Y}_{t-l}(\tau_1)-\overline{Y}_{t-\tau_j-l}(\tau_1)\\&=\frac{1}{2}[\overline{Y}_t(\tau_j)-\overline{Y}_{t-\tau_j}(\tau_j)].\end{aligned}\tag{540}$$

假定向后差分过程$\{\bar{Y}_t(\tau_1)-\bar{Y}_{t-1}(\tau_1)\}$有零均值也意味着$E\{\bar{W}_{j,t}\}=0$(之所以得到这样的结论是因为我们可以记

$$\begin{aligned}\bar{Y}_{t-l}(\tau_1)-\bar{Y}_{t-\tau_j-l}(\tau_1)=&[\bar{Y}_{t-l}(\tau_1)-\bar{Y}_{t-1-l}(\tau_1)]\\&+[\bar{Y}_{t-1-l}(\tau_1)-\bar{Y}_{t-2-l}(\tau_1)]+\cdots\\&+[\bar{Y}_{t-\tau_j+2-l}(\tau_1)-\bar{Y}_{t-\tau_j+1-l}(\tau_1)]\\&+[\bar{Y}_{t-\tau_j+1-l}(\tau_1)-\bar{Y}_{t-\tau_j-l}(\tau_1)],\end{aligned}$$

方括号里面的每一项都有零均值). 对式(540)两边先平方再取期望就得到要求的结果, 因为$\nu\frac{2}{Y}(\tau_j)=\text{var}\{\bar{W}_{j,t}\}=E\{\bar{W}_{j,t}^2\}$.

练习[345]答案 因为$W_{j,t}$并不受环状影响, 我们可以写

$$W_{j,t}=\sum_{l=0}^{L_j-1}h_{j,l}X_{2^j(t+1)-1-l},$$

具有$W_{j',t'}$的一个类似表达(上面是式(96d)去掉"mod$N$"). 式(345a)立即根据练习[7.2]的(d)部分得到, 因为

$$\text{cov}\{X_{2^j(t+1)-1-l},X_{2^{j'}(t'+1)-1-l'}\}=s_{X,2^j(t+1)-l-2^{j'}(t'+1)+l'}.$$

当$j=j'$和$t'=t+\tau$时, 我们有

$$\text{cov}\{W_{j,t},W_{j,t'}\}=\sum_{l=0}^{L_j-1}\sum_{l'=0}^{L_j-1}h_{j,l}h_{j,l'}s_{X,2^j\tau+l-l'}.$$

我们可以把右边的项认为是对称$L_j\times L_j$阶矩阵所有元素的和, 它的第(l,l')个元素由

$$h_{j,l}h_{j,l'}s_{X,2^j\tau+l-l'}$$

给出. 我们计算这个和通过从$l=0$到L_j-1逐行计算; 然而, 也可以通过逐次计算对角线来计算这个和. 考虑第m个这样的对角线, 其中$m=0,1,-1,\cdots$分别是主对角线, 第一个上对角线和下对角线等的指标. m个对角线的元素满足$l-l'=m$, 因此由$h_{j,l}h_{j,m+l}s_{X,2^j\tau+m}$给出, 其中当$m\geqslant0$时$l=0,\cdots,L_j-m-1$. 通过对这$L_j-m$个元素求和, 再回忆所讨论的矩阵是对称阵便得到式(345b).

练习[348a]答案 将式(267c)中给出的$s_{X,\tau}$的表达式(Δt取1, 所以$f_N=1/2$)代入式(345a)中产生

$$\begin{aligned}\text{cov}\{W_{j,t},W_{j',t'}\}&=\sum_{l=0}^{L_j-1}\sum_{l'=0}^{L_{j'}-1}h_{j,l}h_{j',l'}\int_{-1/2}^{1/2}S_X(f)e^{i2\pi f(2^j(t+1)-l-2^{j'}(t'+1)+l')}df\\&=\int_{-1/2}^{1/2}\left(\sum_{l=0}^{L_j-1}h_{j,l}e^{-i2\pi fl}\right)\left(\sum_{l'=0}^{L_{j'}-1}h_{j',l'}e^{-i2\pi fl'}\right)\\&\quad\times S_X(f)e^{i2\pi f(2^j(t+1)-2^{j'}(t'+1))}df\\&=\int_{-1/2}^{1/2}H_j(f)H_{j'}^*(f)S_X(f)e^{i2\pi f(2^j(t+1)-2^{j'}(t'+1))}df.\end{aligned}$$

由于$\text{cov}\{W_{j,t},W_{j',t'}\}=\text{cov}\{W_{j',t'},W_{j,t}\}$, 我们可以用$j'$和$t'$替换$j$和$t$以获得要求的结果(这样的替换允许我们在一些子序列的表示上去掉一些负号).

练习[348b]答案 如果我们对式(348b)右边作变量代换 $f'=2^j f$，得到

$$\operatorname{cov}\{W_{j,t},W_{j,t+\tau}\}=\frac{1}{2^j}\int_{-2^{j-1}}^{2^{j-1}}\mathrm{e}^{\mathrm{i}2\pi f'\tau}\ \mathcal{H}_j\left(\frac{f'}{2^j}\right)S_X\left(\frac{f'}{2^j}\right)\mathrm{d}f'$$

$$=\frac{1}{2^j}\int_{0}^{2^{j}}\mathrm{e}^{\mathrm{i}2\pi f'\tau}\ \mathcal{H}_j\left(\frac{f'}{2^j}\right)S_X\left(\frac{f'}{2^j}\right)\mathrm{d}f',$$

其中最后一行成立时因为积分是以 2^j 为周期的. 如果我们把区间$[0,\ 2^j]$划分成 2^j 个互不重叠的单位长度区间，得到

$$\operatorname{cov}\{W_{j,t},W_{j,t+\tau}\}=\frac{1}{2^j}\sum_{k=0}^{2^j-1}\int_{k}^{k+1}\mathrm{e}^{\mathrm{i}2\pi f'\tau}\ \mathcal{H}_j\left(\frac{f'}{2^j}\right)S_X\left(\frac{f'}{2^j}\right)\mathrm{d}f'$$

$$=\frac{1}{2^j}\sum_{k=0}^{2^j-1}\int_{0}^{1}\mathrm{e}^{\mathrm{i}2\pi(f+k)\tau}\ \mathcal{H}_j\left(\frac{f+k}{2^j}\right)S_X\left(\frac{f+k}{2^j}\right)\mathrm{d}f$$

$$=\frac{1}{2^j}\int_{0}^{1}\mathrm{e}^{\mathrm{i}2\pi f\tau}\sum_{k=0}^{2^j-1}\mathcal{H}_j\left(\frac{f+k}{2^j}\right)S_X\left(\frac{f+k}{2^j}\right)\mathrm{d}f,$$

其中我们作变量代换 $f=f'-k$ 并且利用 $\mathrm{e}^{\mathrm{i}2\pi(f+k)\tau}=\mathrm{e}^{\mathrm{i}2\pi f\tau}$，因为 $k\tau$ 是一个整数. 如果我们能够证明

$$A(f)\equiv\sum_{k=0}^{2^j-1}\mathcal{H}_j\left(\frac{f+k}{2^j}\right)S_X\left(\frac{f+k}{2^j}\right)$$

是以单位为周期的周期函数，那么就得到要求的结果. 为了说明它是正确的，注意

$$A(f+1)=\sum_{k=0}^{2^j-1}\mathcal{H}_j\left(\frac{f+1+k}{2^j}\right)S_X\left(\frac{f+1+k}{2^j}\right)$$

$$=\sum_{k=1}^{2^j}\mathcal{H}_j\left(\frac{f+k}{2^j}\right)S_X\left(\frac{f+k}{2^j}\right)$$

$$=\sum_{k=0}^{2^j-1}\mathcal{H}_j\left(\frac{f+k}{2^j}\right)S_X\left(\frac{f+k}{2^j}\right)=A(f),$$

这里最后一行成立是因为，事实上

$$\mathcal{H}_j\left(\frac{f+2^j}{2^j}\right)S_X\left(\frac{f+2^j}{2^j}\right)=\mathcal{H}_j\left(\frac{f}{2^j}\right)S_X\left(\frac{f}{2^j}\right),$$

因为$\mathcal{H}_j(\cdot)$和 $S_X(\cdot)$都有单位周期.

练习[355]答案 用通用的记号，式(302)记为

$$\operatorname{var}\{X_t\}=\operatorname{var}\{\bar{V}_{J,0}\}+\sum_{j=1}^{J}\nu_X^2(\tau_j),$$

其中 $\bar{V}_{J,0}$是极大重叠小波变换尺度系数，它通过 $V_{J,0}/2^{J/2}=\bar{V}_{J,0}$与离散小波变换尺度系数联系起来，得到 $\operatorname{var}\{V_{J,0}\}/N=\operatorname{var}\{\bar{V}_{J,0}\}$；同样，$\nu_X^2(\tau_j)$是极大重叠小波变换小波系数$\{\widetilde{W}_{j,t}\}$的方差，它通过 $W_{j,t}/2^{j/2}=\widetilde{W}_{j,2^j(t+1)-1}$与离散小波变换小波系数$\{W_{j,t}\}$联系起来得到 $\operatorname{var}\{W_{j,t}\}/2^j=\nu_X^2(\tau_j)$.

练习[362a]答案 我们使用这样的事实，如果 A，B 和 C 是方阵，那么 $|ABC|=|BCA|$ 和 $|AB|=|A|\cdot|B|$，因此

$$|\widetilde{\Sigma}_{\mathbf{X}}|=|\mathcal{W}^{\mathrm{T}}\Lambda_N\mathcal{W}|=|\Lambda_N\mathcal{W}\mathcal{W}^{\mathrm{T}}|=|\Lambda_N I_N|=|\Lambda_N|\cdot|I_N|=|\Lambda_N|.$$

因为恒等矩阵的行列式是 1. 由于确定对角矩阵的行列式是由对角线元素相乘，并且由于对角线元素如式(355a)，那么

$$|\Lambda_N|=C_{J+1}\prod_{j=1}^{J}C_j^{N_j}，\quad 因此\ \log(|\Lambda_N|)=\log(C_{J+1})+\sum_{j=1}^{J}N_j\log(C_j),$$

其中，如通常 $N_j\equiv N/2^j$.

练习[362b]答案 首先注意

$$\widetilde{\Sigma}_{\mathbf{X}}^{-1}=(\mathcal{W}^{\mathrm{T}}\Lambda_N\mathcal{W})^{-1}=\mathcal{W}^{-1}\Lambda_N^{-1}(\mathcal{W}^{\mathrm{T}})^{-1}=\mathcal{W}^{\mathrm{T}}\Lambda_N^{-1}\mathcal{W},$$

其中 Λ_N^{-1} 是一个对角矩阵，它的元素是

$$\underbrace{\frac{1}{C_1},\cdots,\frac{1}{C_1}}_{\frac{N}{2}个},\underbrace{\frac{1}{C_2},\cdots,\frac{1}{C_2}}_{\frac{N}{4}个},\cdots,\underbrace{\frac{1}{C_j},\cdots,\frac{1}{C_j}}_{\frac{N}{2^j}个},\cdots,\underbrace{\frac{1}{C_{J-1}},\frac{1}{C_{J-1}}}_{2个},\frac{1}{C_J},\frac{1}{C_{J+1}}.$$

因此

$$\begin{aligned}\mathbf{X}^{\mathrm{T}}\widetilde{\Sigma}_{\mathbf{X}}^{-1}\mathbf{X}&=\mathbf{X}^{\mathrm{T}}\mathcal{W}^{\mathrm{T}}\Lambda_N^{-1}\mathcal{W}\mathbf{X}=\mathbf{W}^{\mathrm{T}}\Lambda_N^{-1}\mathbf{W}\\&=[\mathbf{W}_1^{\mathrm{T}},\mathbf{W}_2^{\mathrm{T}},\cdots,\mathbf{W}_J^{\mathrm{T}},\mathbf{V}_J^{\mathrm{T}}]\begin{bmatrix}\mathbf{W}_1/C_1\\\mathbf{W}_2/C_2\\\vdots\\\mathbf{W}_J/C_J\\\mathbf{V}_J/C_{J+1}\end{bmatrix}\\&=\frac{\mathbf{V}_J^{\mathrm{T}}\mathbf{V}_J}{C_{J+1}}+\sum_{j=1}^{J}\frac{\mathbf{W}_j^{\mathrm{T}}\mathbf{W}_j}{C_j}\\&=\frac{V_{J,0}^2}{\sigma_\varepsilon^2C'_{J+1}(\delta)}+\sum_{j=1}^{J}\frac{1}{\sigma_\varepsilon^2C'_j(\delta)}\sum_{t=0}^{N_j-1}W_{j,t}^2,\end{aligned}$$

它析出因子 $1/\sigma_\varepsilon^2$ 就得到要求的结果.

练习[365a]答案 由于小波滤波器中的嵌入差分运算，小波系数 $W_{j,t}$ 与时间序列的样本均值无关；另一方面，单个尺度系数 $V_{J,0}$ 明显与样本均值有关，因为，如果我们对一个有着样本均值 $\overline{X}$ 的时间序列 X 进行离散小波变换，那么 $V_{J,0}=\overline{X}\sqrt{N}$. 通过注意时间序列 $\mathbf{X}-\mu\mathbf{1}$ 有一个样本均值 $\overline{X}-\mu$，因此，相应的尺度系数采用形式 $V_{J,0}=(\overline{X}-\mu)\sqrt{N}$，这样就可以立即从练习[362b]的结果中得到所要求的结果.

练习[365b]答案 时间序列 $\mathbf{X}'\equiv\mathbf{X}-\overline{X}\mathbf{1}$ 有一个零样本均值，因此，尽管 $\mathbf{X}$ 和 $\mathbf{X}'$ 的小波系数是相等的，但是尺度系数分别为 $\overline{X}\sqrt{N}$ 和零. 当 $V_{J,0}$ 取零时，式(363b)的极大似然估计 $\hat{\sigma}_\varepsilon^2(\delta)$ 减小至式(365)中的极大似然估计，当过程均值未知时反转而导致对于 δ 的准确的减少的对数似然性.

练习[371]答案 （改变尺度和重新定位的）对数似然函数是

$$\tilde{l}(\delta,\sigma_\varepsilon^2 \mid \mathbf{W}'_{nb}) \equiv -2\log(\widetilde{L}(\delta,\sigma_\varepsilon^2 \mid \mathbf{W}'_{nb})) - N'\log(2\pi)$$
$$= \sum_{j=1}^{J_0}\sum_{t=0}^{N'_j-1}\left(\frac{W^2_{j,t+L'_j}}{C_j} + \log(C_j)\right).$$

回忆 $C_j = \sigma_\varepsilon^2 C'_j(\delta)$，我们得到

$$\tilde{l}(\delta,\sigma_\varepsilon^2 \mid \mathbf{W}'_{nb}) = \frac{1}{\sigma_\varepsilon^2}\sum_{j=1}^{J_0}\frac{1}{C'_j(\delta)}\sum_{t=0}^{N'_j-1}W^2_{j,t+L'_j} + N'\log(\sigma_\varepsilon^2) + \sum_{j=1}^{J_0}N'_j\log(C'_j(\delta)).$$

对右边关于 σ_ε^2 微分，令得到的表达式为零就得到式(371a)中的$\tilde{\sigma}_\varepsilon^2(\delta)$. 把$\tilde{\sigma}_\varepsilon^2(\delta)$代入上式且减掉 N'就得到式(371c)所示的减少的似然函数.

练习[373]答案 从式(370a)，我们有 $\mathbf{X}-\hat{\mathbf{T}}=\widetilde{\mathbf{U}}$，其中$\widetilde{\mathbf{U}} \equiv \mathcal{W}^{\mathrm{T}}\mathbf{W}_{nb}$，因此

$$\|\mathbf{X}-\hat{\mathbf{T}}\|^2 = \|\mathcal{W}^{\mathrm{T}}\mathbf{W}_{nb}\|^2 = \|\mathbf{W}_{nb}\|^2,$$

因为$\mathcal{W}^{\mathrm{T}}$ 是一个规范正交变换.

练习[395]答案 规范正交变换的保持能量性质是说 $\|\mathbf{O}\|^2 = \|\mathbf{D}\|^2$. 因为

$$\|\mathbf{O}\|^2 = \|\mathcal{I}_M\mathbf{O}\|^2 + \|(I_N-\mathcal{I}_M)\mathbf{O}\|^2$$

和

$$\|\mathbf{D}-\hat{\mathbf{D}}_M\|^2 = \|\mathcal{O}^{\mathrm{T}}\mathbf{O}-\mathcal{O}^{\mathrm{T}}\mathcal{I}_M\mathbf{O}\|^2$$
$$= \|\mathcal{O}^{\mathrm{T}}(I_N-\mathcal{I}_M)\mathbf{O}\|^2$$
$$= \|(I_N-\mathcal{I}_M)\mathbf{O}\|^2,$$

我们就得到要求的

$$C_{M-1} = \frac{\|\mathcal{I}_M\mathbf{O}\|^2}{\|\mathbf{O}\|^2}$$
$$= \frac{\|\mathbf{O}\|^2 - \|(I_N-\mathcal{I}_M)\mathbf{O}\|^2}{\|\mathbf{O}\|^2}$$
$$= 1 - \frac{\|\mathbf{D}-\hat{\mathbf{D}}_M\|^2}{\|\mathbf{D}\|^2}.$$

练习[404]答案 注意到

$$A^{(\delta)}(O_l) = -O_l 1_{[0,\delta^2)}(O_l^2) - \delta\,\mathrm{sign}\{O_l\}1_{[\delta^2,\infty)}(O_l^2).$$

因此就得到要求的

$$\mathcal{R}(\sigma_{nl},O_l,\delta) = \sigma_{nl}^2 - 2\sigma_{nl}^2 1_{[0,\delta^2)}(O_l^2) + O_l^2 1_{[0,\delta^2)}(O_l^2) + \delta^2 1_{[\delta^2,\infty)}(O_l^2)$$
$$= -\sigma_{nl}^2 1_{[0,\delta^2)}(O_l^2) + \sigma_{nl}^2 1_{[\delta^2,\infty)}(O_l^2) + O_l^2 1_{[0,\delta^2)}(O_l^2) + \delta^2 1_{[\delta^2,\infty)}(O_l^2)$$
$$= [O_l^2 - \sigma_{nl}^2]1_{[0,\delta^2)}(O_l^2) + [\sigma_{nl}^2 + \delta^2]1_{[\delta^2,\infty)}(O_l^2)$$
$$= O_l^2 - \sigma_{nl}^2 - [O_l^2 - \sigma_{nl}^2]1_{[\delta^2,\infty)}(O_l^2) + [\sigma_{nl}^2 + \delta^2]1_{[\delta^2,\infty)}(O_l^2)$$
$$= O_l^2 - \sigma_{nl}^2 + (2\sigma_{nl}^2 - O_l^2 + \delta^2)1_{[\delta^2,\infty)}(O_l^2).$$

练习[407]答案 我们有

$$E\{(R_l-a_lO_l)^2\}=E\{R_l^2\}-2a_lE\{R_lO_l\}+a_l^2E\{O_l^2\}\equiv f(a_l),$$

它是关于 a_l 的一个二次函数(假定 $E\{O_l^2\}\neq 0$). 为了确定这个函数的极值，我们对上式关于 a_l 微分，令得到的表达式等于零，得到

$$\frac{\mathrm{d}f(a_l)}{\mathrm{d}a_l}=-2E\{R_lO_l\}+2a_lE\{O_l^2\}=0$$

(由于二次导数是 $2E\{O_l^2\}$，因此是正的，如果 $E\{O_l^2\}\neq 0$，极值实际上是极小值). 上式等价于所要求的

$$E\{R_lO_l\}-a_lE\{O_l^2\}=E\{(R_l-a_lO_l)O_l\}=0.$$

练习[441]答案 让 $\boldsymbol{b}^{\mathrm{T}}$ 表示 $\mathcal{W}$ 的行向量，那么在 $\boldsymbol{n}=\mathcal{W}\boldsymbol{\eta}$ 处产生系数 $n_{j,t}$；即 $n_{j,t}=\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\eta}$. 在式(262c)中让 $\mathcal{M}=\boldsymbol{b}^{\mathrm{T}}$，利用前边练习中的对 Σ_η 进行分解并且回忆 $\boldsymbol{b}^{\mathrm{T}}=\boldsymbol{b}^{\mathrm{H}}$，因为 $\mathcal{W}$ 是实值，那么我们有

$$\operatorname{var}\{n_{j,t}\}=\boldsymbol{b}^{\mathrm{T}}\Sigma_\eta\boldsymbol{b}=\boldsymbol{b}^{\mathrm{H}}\mathcal{F}^{\mathrm{H}}D\mathcal{F}\boldsymbol{b}=(\mathcal{F}\boldsymbol{b})^{\mathrm{H}}D\mathcal{F}\boldsymbol{b}.$$

现在 $\mathcal{F}\boldsymbol{b}$ 是 $\boldsymbol{b}$ 的规范正交离散傅里叶变换，它通过 $\mathcal{F}\boldsymbol{b}=\frac{1}{\sqrt{M}}[B_0B_1,\cdots,B_{M-1}]^{\mathrm{T}}$ 与 $\boldsymbol{b}$ 的傅里叶变换 $\{B_k\}$ 联系起来(比较式(46b)和(36g)). 因此

$$\begin{aligned}\operatorname{var}\{n_{j,t}\}&=\frac{1}{M}[B_0^*,B_1^*,\cdots,B_{M-1}^*]\begin{bmatrix}S_0&0&\cdots&0\\0&S_1&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&S_{M-1}\end{bmatrix}\begin{bmatrix}B_0\\B_1\\\vdots\\B_{M-1}\end{bmatrix}\\&=\frac{1}{M}[B_0^*,B_1^*,\cdots,B_{M-1}^*]\begin{bmatrix}S_0B_0\\S_1B_1\\\vdots\\S_{M-1}B_{M-1}\end{bmatrix}\\&=\frac{1}{M}\sum_{k=0}^{M-1}S_k|B_k|^2.\end{aligned}$$

注意 $|B_k|^2=\left|H_j\left(\frac{k}{N}\right)\right|^2=\mathcal{H}_j\left(\frac{k}{N}\right)$ 得到式(441b).

练习[464]答案 让 $u\equiv 2^{-j}t-2k$，所以

$$2^{-j}t-m=u+2k-m,\qquad 2^{-(j+1)}t-k=2^{-1}u \text{ 和 } 2^{-j}\mathrm{d}t=\mathrm{d}u. \tag{545}$$

那么

$$\begin{aligned}\langle\phi_{j,m}(\cdot),\phi_{j+1,k}(\cdot)\rangle&\equiv\int_{-\infty}^{\infty}\phi_{j,m}(t)\phi_{j+1,k}(t)\mathrm{d}t\\&=\int_{-\infty}^{\infty}2^{-j}\phi(2^{-j}t-m)2^{-1/2}\phi(2^{-(j+1)}t-k)\mathrm{d}t\\&=\int_{-\infty}^{\infty}\phi(u+2k-m)2^{-1/2}\phi(2^{-1}u)\mathrm{d}u\\&=\int_{-\infty}^{\infty}\phi_{0,m-2k}(u)\phi_{1,0}(u)\mathrm{d}u=\langle\phi_{0,m-2k}(\cdot),\phi_{1,0}(\cdot)\rangle=\bar{g}_{m-2k}.\end{aligned}$$

为了证明推论是正确的，注意，由于 $j'\equiv j-1$，$m'=k$ 和 $k'=m$，

$$\langle\phi_{j,m}(\cdot),\phi_{j-1,k}(\cdot)\rangle=\langle\phi_{j-1,k}(\cdot),\phi_{j,m}(\cdot)\rangle=\langle\phi_{j',m'}(\cdot),\phi_{j'+1,k'}(\cdot)\rangle=\bar{g}_{m'-2k'}=\bar{g}_{k-2m},$$

其中我们利用了第一部分.

练习[466]答案 对式(466c)两边进行傅里叶变换得到要求的

$$\begin{aligned}\Phi_{j+1}(f) &= \sum_{l=-\infty}^{\infty}\bar{g}_l\int_{-\infty}^{\infty}\phi_{j,0}(t-2^jl)\mathrm{e}^{-i2\pi ft}\,\mathrm{d}t\\ &= \sum_{l=-\infty}^{\infty}\bar{g}_l\mathrm{e}^{-i2\pi 2^jfl}\int_{-\infty}^{\infty}\phi_{j,0}(t-2^jl)\mathrm{e}^{-i2\pi f(t-2^jl)}\,\mathrm{d}t\\ &= \bar{G}(2^jf)\int_{-\infty}^{\infty}\phi_{j,0}(t')\mathrm{e}^{-i2\pi ft'}\,\mathrm{d}t'\\ &= \bar{G}(2^jf)\Phi_j(f).\end{aligned}$$

练习[468a]答案 首先我们说明，任何非零的整数 m 可以写成 $m=2^l[2j+1]$，对于一些 $l\geqslant 0$ 和 $j\in \mathbf{Z}$. 注意，为了验证它设 $l=0$，我们得到 $m=2j+1$，它产生所有的奇数；另一方面，我们可以把任意偶数 m 看作一些奇数 $2j+1$ 和一些 2^l 而 $l\geqslant 1$ 的乘积. 在式(465c)中置 $n=l+1$，我们得到

$$\begin{aligned}\Phi(m) &= \Phi\left(\frac{2^l[2j+1]}{2^{l+1}}\right)\prod_{k=1}^{l+1}\frac{\bar{G}\left(\frac{2^l[2j+1]}{2^k}\right)}{\sqrt{2}}\\ &= \Phi\left(\frac{2j+1}{2}\right)\prod_{k=1}^{l+1}\frac{\bar{G}(2^{l-k}[2j+1])}{\sqrt{2}}.\end{aligned}$$

乘积中的 $k=l-1$ 项包含 $\bar{G}\left(j+\frac{1}{2}\right)$；然而，$\bar{G}(\cdot)$是周期为 1 的周期函数，因此从式(468a)有$G\left(j+\frac{1}{2}\right)=G\left(\frac{1}{2}\right)=0$. 所以 $\Phi(m)=0$(这里用到的方法是基于 Daubechies，1992，p. 174-175).

练习[468b]答案 首先注意，式(465a)的两尺度关系推出

$$\phi(t+k)=\sqrt{2}\sum_{l'=-\infty}^{\infty}\bar{g}_{l'}\phi(2t+2k-l'),$$

因此

$$\phi(t)\phi(t+k)=2\sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\bar{g}_l\bar{g}_{l'}\phi(2t-l)\phi(2t+2k-l').$$

两边对 t 进行积分，得到所要求的

$$\begin{aligned}\delta_{k,0} &= 2\sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\bar{g}_l\bar{g}_{l'}\int_{-\infty}^{\infty}\phi(2t-l)\phi(2t+2k-l')\mathrm{d}t\\ &= \sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\bar{g}_l\bar{g}_{l'}\int_{-\infty}^{\infty}\phi(t-l)\phi(t+2k-l')\mathrm{d}t\\ &= \sum_{l=-\infty}^{\infty}\bar{g}_l\bar{g}_{l+2k}.\end{aligned}$$

练习[474a]答案

$$\int_{-\infty}^{\infty}\psi_{0,m}(t)\phi_{0,n}(t)\,\mathrm{d}t$$

$$=\int_{-\infty}^{\infty}\psi(t-m)\phi(t-n)\,\mathrm{d}t$$

$$=2\int_{-\infty}^{\infty}\left[\sum_{l=-\infty}^{\infty}\bar{h}_l\phi(2t-2m-l)\right]\left[\sum_{l'=-\infty}^{\infty}\bar{g}_{l'}\phi(2t-2n-l')\right]\mathrm{d}t$$

$$=2\sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\bar{h}_l\bar{g}_{l'}\int_{-\infty}^{\infty}\phi(2t-2m-l)\phi(2t-2n-l')\,\mathrm{d}t$$

$$=\sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\bar{h}_l\bar{g}_{l'}\int_{-\infty}^{\infty}\phi(t-2m-l)\phi(t-2n-l')\,\mathrm{d}t$$

$$=\sum_{l=-\infty}^{\infty}\bar{h}_l\bar{g}_{l+2m-2n}$$

$$=\sum_{l=-\infty}^{\infty}\bar{h}_{l+n'}\bar{g}_{l-n'}\qquad\text{置 } n'\equiv n-m$$

$$=\sum_{l=-\infty}^{\infty}(-1)^{l+n'}\bar{g}_{1-l-n'-L}\bar{g}_{l-n'}$$

$$=\sum_{l=-\infty}^{-L/2}(-1)^{l+n'}\bar{g}_{1-l-n'-L}\bar{g}_{l-n'}+\sum_{l=-(L/2)+1}^{\infty}(-1)^{l+n'}\bar{g}_{1-l-n'-L}\bar{g}_{l-n'}$$

$$=\sum_{l=L/2}^{\infty}(-1)^{-l+n'}\bar{g}_{1+l-n'-L}\bar{g}_{-l-n'}-\sum_{l=L/2}^{\infty}(-1)^{l+n'}\bar{g}_{-l-n'}\bar{g}_{1+l-n'-L}$$

$$=0,$$

其中最后一行是由于对于所有的 l 和 n' 且因为 L 是偶数，$(-1)^{-l+n'}=(-1)^{l+n'-L}=(-1)^{l+n'}$. 另一个可以替代的证明是注意通过引用练习[77]的结果

$$\sum_{l=-\infty}^{\infty}\bar{h}_l\bar{g}_{l+2m-2n}=\sum_{l=-\infty}^{\infty}h_{-l}g_{-l+2n-2m}$$

$$=\sum_{l=-\infty}^{\infty}h_l g_{l+2(n-m)}=0.$$

练习[474b]答案 我们有

$$\int_{-\infty}^{\infty}\psi_{0,m}(t)\psi_{0,n}(t)\,\mathrm{d}t=\int_{-\infty}^{\infty}\psi(t-m)\psi(t-n)\,\mathrm{d}t$$

$$=2\int_{-\infty}^{\infty}\left[\sum_{l=-\infty}^{\infty}\bar{h}_l\phi(2t-2m-l)\right]\left[\sum_{l'=-\infty}^{\infty}\bar{h}_{l'}\phi(2t-2n-l')\right]\mathrm{d}t$$

$$=2\sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\bar{h}_l\bar{h}_{l'}\int_{-\infty}^{\infty}\phi(2t-2m-l)\phi(2t-2n-l')\,\mathrm{d}t$$

$$= \sum_{l=-\infty}^{\infty}\sum_{l'=-\infty}^{\infty}\bar{h}_l\bar{h}_{l'}\int_{-\infty}^{\infty}\phi(t-2m-l)\phi(t-2n-l')\mathrm{d}t$$

$$= \sum_{l=-\infty}^{\infty}\bar{h}_l\bar{h}_{l+2m-2n}$$

$$= \sum_{l=-\infty}^{\infty}(-1)^{2l+2m-2n}\bar{g}_{1-l}\bar{g}_{1-l-2m+2n}$$

$$= \sum_{l=-\infty}^{\infty}\bar{g}_{1-l}\bar{g}_{1-l-2m+2n}$$

$$= \sum_{l=-\infty}^{\infty}g_{l-1}g_{l-1-2n+2m}$$

$$= \sum_{l=-\infty}^{\infty}g_lg_{l+2(m-n)} = \delta_{m,n},$$

因为单位能量滤波器$\{g_l\}$与它的非零偶平移是正交的(参见练习[76b]).

练习[476]答案 由于$\bar{h}_l\equiv(-1)^l\bar{g}_{1-l-L}$(式(472)),我们有(令 $m\equiv 1-l-L$)

$$\bar{H}(f)\equiv\sum_{l=-\infty}^{\infty}\bar{h}_l\mathrm{e}^{-\mathrm{i}2\pi fl}=\sum_{l=-(L-1)}^{0}(-1)^l\bar{g}_{1-l-L}\mathrm{e}^{-\mathrm{i}2\pi fl}$$

$$=\sum_{m=-(L-1)}^{0}(-1)^{1-m-L}\bar{g}_m\mathrm{e}^{-\mathrm{i}2\pi f(1-m-L)}$$

$$=-\mathrm{e}^{-\mathrm{i}2\pi f(1-L)}\sum_{m=-(L-1)}^{0}(-1)^{-m-L}\bar{g}_m\mathrm{e}^{\mathrm{i}2\pi fm}$$

$$=-\mathrm{e}^{-\mathrm{i}2\pi f(1-L)}\sum_{m=-(L-1)}^{0}(-1)^{-L}\mathrm{e}^{-\mathrm{i}\pi m}\bar{g}_m\mathrm{e}^{\mathrm{i}2\pi fm}$$

$$=-\mathrm{e}^{-\mathrm{i}2\pi f(1-L)}\sum_{m=-(L-1)}^{0}\bar{g}_m\mathrm{e}^{\mathrm{i}2\pi\left(f-\frac{1}{2}\right)m}$$

$$=-\mathrm{e}^{-\mathrm{i}2\pi f(1-L)}\left(\sum_{m=-(L-1)}^{0}\bar{g}_m\mathrm{e}^{-\mathrm{i}2\pi\left(f-\frac{1}{2}\right)m}\right)^*$$

$$=-\mathrm{e}^{\mathrm{i}2\pi f(L-1)}\bar{G}^*\left(f-\frac{1}{2}\right)$$

$$=-\mathrm{e}^{\mathrm{i}2\pi f(L-1)}\bar{G}\left(\frac{1}{2}-f\right)$$

(我们利用了当 L 是偶数时$(-1)^{-L}=1$ 这一事实).

练习[479]答案 我们有

$$\langle\phi_{0,m-2k}(\cdot),\phi_{1,0}(\cdot)\rangle=\int_{-\infty}^{\infty}\phi(t-[m-2k])\frac{\phi\left(\frac{t}{2}\right)}{\sqrt{2}}\mathrm{d}t$$

$$=\int_{-\infty}^{\infty}\phi(t-[m-2k])\left[\sum_{l=-\infty}^{\infty}\bar{h}_l\phi(t-l)\right]\mathrm{d}t$$

$$=\sum_{l=-\infty}^{\infty}\bar{h}_l\int_{-\infty}^{\infty}\phi(t-[m-2k])\phi(t-l)\mathrm{d}t$$

$$=\sum_{l=-\infty}^{\infty}\bar{h}_l\delta_{l,m-2k}=\bar{h}_{m-2k}.$$

如像练习[464]中，让 $u=2^{-j}t-2k$，因此式(545)成立．那么，就得到要求的

$$\begin{aligned}\langle \phi_{j,m}(\cdot),\psi_{j+1,k}(\cdot)\rangle &= \int_{-\infty}^{\infty} \phi_{j,m}(t)\psi_{j+1,k}(t)\mathrm{d}t \\ &= \int_{-\infty}^{\infty} 2^{-j}\phi(2^{-j}t-m)2^{-1/2}\psi(2^{-(j+1)}t-k)\mathrm{d}t \\ &= \int_{-\infty}^{\infty} \phi(u+2k-m)2^{-1/2}\psi(2^{-1}u)\mathrm{d}u \\ &= \int_{-\infty}^{\infty} \phi_{0,m-2k}(u)\psi_{1,0}(u)\mathrm{d}u \\ &= \langle \phi_{0,m-2k}(\cdot),\psi_{1,0}(\cdot)\rangle = \bar{h}_{m-2k}.\end{aligned}$$

练习[481]答案 我们有

$$\begin{aligned}v_{2,k} &= \langle s_1(\cdot),\phi_{2,k}(\cdot)\rangle \\ &= \langle x(\cdot)-d_1(\cdot),\phi_{2,k}(\cdot)\rangle \\ &= \langle x(\cdot),\phi_{2,k}(\cdot)\rangle - \langle d_1(\cdot),\phi_{2,k}(\cdot)\rangle \\ &= \langle x(\cdot),\phi_{2,k}(\cdot)\rangle - \sum_{m=-\infty}^{\infty} w_{1,m}\langle \psi_{1,m}(\cdot),\phi_{2,k}(\cdot)\rangle \\ &= \langle x(\cdot),\phi_{2,k}(\cdot)\rangle,\end{aligned}$$

由于

$$\phi_{j,k}(\cdot) \perp \psi_{l,m}(\cdot) \quad \text{对于 } j,k,l,m \in \mathbf{Z} \text{ 因此 } l \leqslant j.$$

类似地，

$$\begin{aligned}w_{2,k} &= \langle s_1(\cdot),\psi_{2,k}(\cdot)\rangle \\ &= \langle x(\cdot)-d_1(\cdot),\psi_{2,k}(\cdot)\rangle \\ &= \langle x(\cdot),\psi_{2,k}(\cdot)\rangle - \langle d_1(\cdot),\psi_{2,k}(\cdot)\rangle \\ &= \langle x(\cdot),\psi_{2,k}(\cdot)\rangle - \sum_{m=-\infty}^{\infty} w_{1,m}\langle \psi_{1,m}(\cdot),\psi_{2,k}(\cdot)\rangle \\ &= \langle x(\cdot),\psi_{2,k}(\cdot)\rangle,\end{aligned}$$

由于

$$\psi_{j,k}(\cdot) \perp \psi_{l,m}(\cdot) \quad \text{对于 } j,k,l,m \in \mathbf{Z} \text{ 因此 } j \neq l.$$

对于一般的情形注意 $x(t)=s_{j-1}(t)+d_{j-1}(t)+\cdots+d_1(t)$．因此，

$$\begin{aligned}v_{j,k} &= \langle s_{j-1}(\cdot),\phi_{j,k}(\cdot)\rangle \\ &= \langle x(\cdot)-\sum_{l=1}^{j-1} d_l(\cdot),\phi_{j,k}(\cdot)\rangle \\ &= \langle x(\cdot),\phi_{j,k}(\cdot)\rangle - \sum_{l=1}^{j-1}\langle d_l(\cdot),\phi_{j,k}(\cdot)\rangle \\ &= \langle x(\cdot),\phi_{j,k}(\cdot)\rangle - \sum_{l=1}^{j-1}\sum_{m=-\infty}^{\infty} w_{l,m}\langle \psi_{l,m}(\cdot),\phi_{j,k}(\cdot)\rangle \\ &= \langle x(\cdot),\phi_{j,k}(\cdot)\rangle,\end{aligned}$$

这里我们又一次用到了

$$\phi_{j,k}(\cdot) \perp \psi_{l,m}(\cdot) \quad \text{对于 } j,k,l,m \in \mathbf{Z} \text{ 因此 } l \leqslant j.$$

对于 $w_{j,k}$ 一般情况类似地得到．

练习[482]答案 从多分辨分析的定义，我们有对任意的 j，有 $V_{j+1}\subset V_j$，它意味着

$$\bigcup_{l\in\mathbf{Z}} V_l = \bigcup_{-\infty<l\le j} V_l.$$

另一方面，事实 $V_{j-1}=V_j\oplus W_j$ 意味着

$$\begin{aligned}V_j \cup V_{j-1} \cup V_{j-2} &= V_j \cup [V_j \oplus W_j] \cup [V_{j-1} \oplus W_{j-1}]\\ &= V_j \cup [V_j \oplus W_j] \cup [V_j \oplus W_j \oplus W_{j-1}].\end{aligned}$$

由直和的定义，对于任意的集合 A 和 B，我们有 $A\subset A\oplus B$，转而它又暗含了 $A\cup[A\oplus B]=A\oplus B$. 因此

$$V_j \cup V_{j-1} \cup V_{j-2} = V_j \oplus W_j \oplus W_{j-1}.$$

因此我们可以证明所要求的

$$\begin{aligned}L^2(\mathbf{R}) &= \overline{\bigcup_{l\in\mathbf{Z}} V_l}\\ &= \overline{V_j \cup V_{j-1} \cup V_{j-2} \cup \cdots}\\ &= \overline{V_j \oplus W_j \oplus W_{j-1} \oplus \cdots}.\end{aligned}$$

练习[484]答案 我们希望证明，对于 $m=0,\cdots,n$，$\Psi^{(m)}(0)=0$ 意味着，对于 $m=0,\cdots,n$，$\overline{H}^{(m)}(0)=0$. 我们使用归纳法. 由于 $\Psi(f)=\Phi\left(\frac{f}{2}\right)\overline{H}\left(\frac{f}{2}\right)/\sqrt{2}$，那么结果对于 $n=0$ 和 1 时成立的，因此 $\Psi(0)=0$ 意味着 $\overline{H}(0)=0$，($\Phi(0)=1$)，我们仅需要证明 $\Psi^{(1)}(0)$推出 $H^{(1)}(0)=0$. 假设结果对于 $n-1$ 成立. 现在假定对于 $m=0,\cdots,n$，$\Psi^{(m)}(0)=0$. 从对乘积微分的莱布尼茨(Leibniz)定理知

$$\Psi^{(n)}(f) = \frac{1}{\sqrt{2}}\sum_{m=0}^{n}\binom{n}{m}\Phi^{(n-m)}\left(\frac{f}{2}\right)\overline{H}^{(m)}\left(\frac{f}{2}\right);$$

但是，如果对于 $m=0,\cdots,n$，$\Psi^{(m)}(0)=0$，那么对于 $m=0,\cdots,n-1$，$\Psi^{(m)}(0)=0$. 由于我们假设结果对于 $n-1$ 成立，我们对于 $m=0,\cdots,n-1$，$H^{(m)}(0)=0$；然而，由莱布尼茨公式

$$\Psi^{(n)}(0) = \frac{1}{\sqrt{2}}\left[\sum_{m=0}^{n-1}\binom{n}{m}\Phi^{(n-m)}(0)\overline{H}^{(m)}(0) + \Phi(0)\overline{H}^{(n)}(0)\right].$$

因此 $\Psi^{(n)}(0)=0=\Phi(0)H^{(n)}(0)$. 由于 $\Phi(0)=1$，我们有 $\overline{H}^{(n)}(0)=0$，它和对于 $m=0,\cdots,n-1$, $\overline{H}^{(m)}(0)=0$ 一起表示对于 $m=0,\cdots,n$，$\overline{H}^{(m)}(0)=0$，结论得证.

练习[485]答案 我们有$\{\overline{g}_l\}\leftrightarrow\overline{G}(\cdot)$，它意味着$\{\overline{g}\star\overline{g}_l\}\leftrightarrow\overline{\mathcal{G}}(\cdot)$，其中$\{\overline{g}\star\overline{g}_l\}$是$\{\overline{g}_l\}$的自相关(见式(25b)和(25c)). 因为$\overline{g}_l$ 是实值的，$\{\overline{g}\star\overline{g}_l\}$必须也是实值的，因此我们有

$$\begin{aligned}\overline{g}\star\overline{g}_l &= (\overline{g}\star\overline{g}_l)^*\\ &= \left(\int_{-1/2}^{1/2}\overline{\mathcal{G}}(\cdot)e^{i2\pi fl}df\right)^*\\ &= \int_{-1/2}^{1/2}\overline{\mathcal{G}}(\cdot)e^{-i2\pi fl}df = \overline{g}\star\overline{g}_{-l},\end{aligned}$$

因此$\{\bar{g}\star\bar{g}_l\}$是一个偶序列．所以

$$\bar{\mathcal{G}}(f)=\sum_{l=-(L-1)}^{L-1}\bar{g}\star\bar{g}_l\mathrm{e}^{-\mathrm{i}2\pi fl}=\bar{g}\star\bar{g}_0+2\sum_{l=1}^{L-1}\bar{g}\star\bar{g}_l\cos(2\pi fl).$$

关系$\cos(2\pi fl)=2\cos(2\pi f[l-1])\cos(2\pi f)-\cos(2\pi f[l-2])$可以重复用，把上式的求和简化为$\cos(2\pi f)$幂的一个线性组合．最后，因为

$$\bar{\mathcal{G}}(f)=\cos^{2r}(\pi f)\,\mathcal{Q}(f)=\left(\frac{1+\cos(2\pi f)}{2}\right)^r\mathcal{Q}(f),$$

我们可以断定$\mathcal{Q}(f)$必定还是一个关于$\cos(2\pi f)$的多项式．

练习[488]答案 首先注意，由于$z=\mathrm{e}^{\mathrm{i}2\pi f}$，由此得出$z^*=z^{-1}$，因为$|z|^2=zz^*=1$．因此

$$[\bar{\bar{Q}}_0(z)]^*=\prod_{l=1}^{\frac{L}{2}-1}(z^{-1}-z_l^*)=\prod_{l=1}^{\frac{L}{2}-1}(z^{-1}-z_l)=\bar{\bar{Q}}_0(z^{-1}),$$

其中中间的等号成立，因为，为了使$\bar{\bar{Q}}_0(z)$是一个实值滤波器z变换，根z_l必须要么实值(在这种情况下$z_l^*=z_l$)或者成对出现，彼此之间互为复共轭．

练习[490]答案 展开$\bar{\bar{G}}(z)$，我们得到

$$\bar{\bar{G}}(z)=\frac{1-\sqrt{3}}{4\sqrt{2}}(b_1z^3+[1+2b_1]z^2+[2+b_1]z+1);$$

然而，比较上式和式(490b)，发现多项式的系数现在颠倒了，由此得出上式导致一个D(4)尺度滤波器的逆滤波器．

参考文献

Abramovich, F., Sapatinas, T. and Silverman, B. W. (1998) Wavelet Thresholding via a Bayesian Approach. *Journal of the Royal Statistical Society, Series B*, 60, 725–49.

Abramowitz, M. and Stegun, I. A., editors (1964) *Handbook of Mathematical Functions*. Washington, DC: US Government Printing Office (reprinted in 1968 by Dover Publications, New York).

Abry, P. and Flandrin, P. (1994) On the Initialization of the Discrete Wavelet Transform Algorithm. *IEEE Signal Processing Letters*, 1, 32–4.

Abry, P., Gonçalvès, P. and Flandrin, P. (1993) Wavelet-Based Spectral Analysis of $1/f$ Processes. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Minneapolis, 3, 237–40.

——— (1995) Wavelets, Spectrum Analysis and $1/f$ Processes. In *Wavelets and Statistics* (Lecture Notes in Statistics, Volume 103), edited by A. Antoniadis and G. Oppenheim. New York: Springer–Verlag, 15–29.

Abry, P. and Veitch, D. (1998) Wavelet Analysis of Long-Range-Dependent Traffic. *IEEE Transactions on Information Theory*, 44, 2–15.

Allan, D. W. (1966) Statistics of Atomic Frequency Standards. *Proceedings of the IEEE*, 54, 221–30.

Anderson, T. W. (1971) *The Statistical Analysis of Time Series*. New York: John Wiley & Sons.

Balek, J. (1977) *Hydrology and Water Resources in Tropical Africa*, Volume 8 of *Developments in Water Science*. New York: Elsevier Scientific Publishing Company.

Baliunas, S., Frick, P., Sokoloff, D. and Soon, W. (1997) Time Scales and Trends in the Central England Temperature Data (1659–1990): A Wavelet Analysis. *Geophysical Research Letters*, 24, 1351–4.

Balogh, A., Smith, E. J., Tsurutani, B. T., Southwood, D. J., Forsyth, R. J. and Horbury, T. S. (1995) The Heliospheric Magnetic Field over the South Polar Region of the Sun. *Science*, 268, 1007–10.

Barnes, J. A., Chi, A. R., Cutler, L. S., Healey, D. J., Leeson, D. B., McGunigal, T. E., Mullen, J. A., Jr., Smith, W. L., Sydnor, R. L., Vessot, R. F. C. and Winkler, G. M. R. (1971) Characterization of Frequency Stability. *IEEE Transactions on Instrumentation and Measurement*, 20, 105–20.

Bartlett, M. S. and Kendall, D. G. (1946) The Statistical Analysis of Variance-Heterogeneity and the Logarithmic Transformation. *Supplement to the Journal of the Royal Statistical Society*, 8, 128–38.

Beauchamp, K. G. (1984) *Applications of Walsh and Related Functions*. London: Academic Press.

Beran, J. (1994) *Statistics for Long-Memory Processes*. New York: Chapman & Hall.

Beran, J. and Terrin, N. (1996) Testing for a Change of the Long-Memory Parameter. *Biometrika*, 83, 627–38.

Best, D. J. and Roberts, D. E. (1975) The Percentage Points of the χ^2 Distribution: Algorithm AS 91. *Applied Statistics*, 24, 385–8.

Beylkin, G. (1992) On the Representation of Operators in Bases of Compactly Supported Wavelets. *SIAM Journal on Numerical Analysis*, 29, 1716–40.

Billingsley, P. (1968) *Convergence of Probability Measures*. New York: John Wiley & Sons.

Blackman, R. B. and Tukey, J. W. (1958) *The Measurement of Power Spectra*. New York: Dover Publications.

Bloomfield, P. (1976) *Fourier Analysis of Time Series: An Introduction*. New York: John Wiley & Sons.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994) *Time Series Analysis: Forecasting and Control* (3rd Edition). Englewood Cliffs, New Jersey: Prentice Hall.

Box, G. E. P. and Tiao, G. C. (1973) *Bayesian Inference in Statistical Analysis*. Reading, Massachusetts: Addison–Wesley.

Bracewell, R. N. (1978) *The Fourier Transform and Its Applications* (Second Edition). New York: McGraw–Hill.

Bradshaw, G. A. and Spies, T. A. (1992) Characterizing Canopy Gap Structure in Forests using Wavelet Analysis. *Journal of Ecology*, 80, 205–15.

Briggs, W. L. and Henson, V. E. (1995) *The DFT: An Owner's Manual for the Discrete Fourier Transform*. Philadelphia: SIAM.

Brillinger, D. R. (1981) *Time Series: Data Analysis and Theory* (Expanded Edition). San Francisco: Holden–Day.

——— (1994) Some River Wavelets. *Environmetrics*, 5, 211–20.

——— (1996) Some Uses of Cumulants in Wavelet Analysis. *Journal of Nonparametric Statistics*, 6, 93–114.

Brockwell, P. J. and Davis, R. A. (1991) *Time Series: Theory and Methods* (Second Edition). New York: Springer.

Brown, J. W. and Churchill, R. V. (1995) *Complex Variables and Applications* (Sixth Edition). New York: McGraw–Hill.

Brown, R. L., Durbin, J. and Evans, J. M. (1975) Techniques for Testing the Constancy of Regression Relationships over Time. *Journal of the Royal Statistical Society, Series B*, 37, 149–63.

Bruce, A. G., Donoho, D. L., Gao, H.-Y. and Martin, R. D. (1994) Denoising and Robust Non-Linear Wavelet Analysis. In *Wavelet Applications* (Proceedings of the SPIE 2242), edited by H. H. Szu. Bellingham, Washington: SPIE Press, 325–36.

Bruce, A. G. and Gao, H.-Y. (1996a) *Applied Wavelet Analysis with S-PLUS*. New York: Springer.

——— (1996b) Understanding WaveShrink: Variance and Bias Estimation. *Biometrika*, 83, 727–45.

Bruce, A. G., Gao, H.-Y. and Stuetzle, W. (1999) Subset-Selection and Ensemble Methods for Wavelet De-Noising. *Statistica Sinica*, 9, 167–82.

Calderón, A. P. (1964) Intermediate Spaces and Interpolation, the Complex Method. *Studia Mathematica*, 24, 113–90.

Carmona, R., Hwang, W.-L. and Torrésani, B. (1998) *Practical Time-Frequency Analysis*. San Diego: Academic Press.

Casella, G. and Berger, R. L. (1990) *Statistical Inference*. Pacific Grove, California: Wadsworth & Brooks/Cole.

Chambers, J. M., Cleveland, W. S., Kleiner, B. and Tukey, P. A. (1983) *Graphical Methods for Data Analysis*. Boston: Duxbury Press.

Chipman, H. A., Kolaczyk, E. D. and McCulloch, R. E. (1997) Adaptive Bayesian Wavelet Shrinkage. *Journal of the American Statistical Association*, 92, 1413–21.

Chui, C. K. (1997) *Wavelets: A Mathematical Tool for Signal Processing*. Philadelphia: SIAM.

Cleveland, W. S. and Parzen, E. (1975) The Estimation of Coherence, Frequency Response, and Envelope Delay. *Technometrics*, 17, 167–72.

Cohen, A., Daubechies, I. and Vial, P. (1993) Wavelets on the Interval and Fast Wavelet Transforms. *Applied and Computational Harmonic Analysis*, 1, 54–81.

Coifman, R. R. and Donoho, D. L. (1995) Translation-Invariant De-Noising. In *Wavelets and Statistics* (Lecture Notes in Statistics, Volume 103), edited by A. Antoniadis and G. Oppenheim. New York: Springer–Verlag, 125–50.

Coifman, R. R., Meyer, Y. and Wickerhauser, M. V. (1992) Size Properties of Wavelet Packets. In *Wavelets and Their Applications*, edited by M. B. Ruskai, G. Beylkin, R. R. Coifman, I. Daubechies, S. Mallat, Y. Meyer and L. Raphael. Boston: Jones and Bartlett, 453–70.

Coifman, R. R. and Wickerhauser, M. V. (1992) Entropy-Based Algorithms for Best Basis Selection. *IEEE Transactions on Information Theory*, 38, 713–8.

Craigmile, P. F., Percival, D. B. and Guttorp, P. (2000a) Wavelet-Based Parameter Estimation for Trend Contaminated Fractionally Differenced Processes. Submitted to *Journal of Time Series Analysis*.

——— (2000b) Trend Assessment Using the Discrete Wavelet Transform. Technical Report, National Research Center for Statistics and the Environment, University of Washington, Seattle.

Crouse, M. S., Nowak, R. D. and Baraniuk, R. G. (1998) Wavelet-Based Statistical Signal Processing Using Hidden Markov Models. *IEEE Transactions on Signal Processing*, 46, 886–902.

Dahlquist, G. and Björck, Å. (1974) *Numerical Methods*. Englewood Cliffs, New Jersey: Prentice Hall.

Daniels, H. E. (1987) Tail Probability Approximations. *International Statistical Review*, 55, 37–48.

Daubechies, I. (1988) Orthonormal Bases of Compactly Supported Wavelets. *Communications on Pure and Applied Mathematics*, 41, 909–96.

——— (1992) *Ten Lectures on Wavelets*. Philadelphia: SIAM.

——— (1993) Orthonormal Bases of Compactly Supported Wavelets II. Variations on a Theme. *SIAM Journal on Mathematical Analysis*, 24, 499–519.

David, H. A. (1985) Bias of S^2 Under Dependence. *The American Statistician*, 39, 201.

Davies, R. B. and Harte, D. S. (1987) Tests for Hurst Effect. *Biometrika*, 74, 95–101.

Davis, G., Mallat, S. G. and Zhang, Z. (1994) Adaptive Time-Frequency Approximations with Matching Pursuits. In *Wavelets: Theory, Algorithms, and Applications*, edited by C. K. Chui, L. Montefusco and L. Puccio. San Diego: Academic Press, 271–93.

Davison, A. C. and Hinkley, D. V. (1997) *Bootstrap Methods and their Application*. Cambridge, UK: Cambridge University Press.

Del Marco, S. and Weiss, J. (1997) Improved Transient Signal Detection Using a Wavepacket-Based Detector with an Extended Translation-Invariant Wavelet Transform. *IEEE Transactions on Signal Processing*, 45, 841–50.

Dijkerman, R. W. and Mazumdar, R. R. (1994) On the Correlation Structure of the Wavelet Coefficients of Fractional Brownian Motion. *IEEE Transactions on Information Theory*, 40, 1609–12.

Donoho, D. L. and Johnstone, I. M. (1994) Ideal Spatial Adaptation by Wavelet Shrinkage. *Biometrika*, 81, 425–55.

Doroslovački, M. I. (1998) On the Least Asymmetric Wavelets. *IEEE Transactions on Signal Processing*, 46, 1125–30.

Draper, N. R. and Smith, H. (1998) *Applied Regression Analysis* (Third Edition). New York: John Wiley & Sons.

Eltahir, E. A. B. and Wang, G. (1999) Nilometers, El Niño, and Climate Variability. *Geophysical Research Letters*, 26, 489–92.

Flandrin, P. (1989) On the Spectrum of Fractional Brownian Motions. *IEEE Transactions on Information Theory*, 35, 197–9.

——— (1992) Wavelet Analysis and Synthesis of Fractional Brownian Motion. *IEEE Transactions on Information Theory*, 38, 910–17.

Foufoula Georgiou, E. and Kumar, P., editors (1994) *Wavelets in Geophysics*. San Diego: Academic Press.

Fuller, W. A. (1996) *Introduction to Statistical Time Series* (Second Edition). New York: John Wiley & Sons.

Gamage, N. K. K. (1990) Detection of Coherent Structures in Shear Induced Turbulence using Wavelet Transform Methods. In *Ninth Symposium on Turbulence and Diffusion*. Boston: American Meteorological Society, 389–92.

Gao, H.-Y. (1993) Wavelet Estimation of Spectral Densities in Time Series Analysis. Ph.D. dissertation, Department of Statistics, University of California, Berkeley.

——— (1997) Choice of Thresholds for Wavelet Shrinkage Estimate of the Spectrum. *Journal of Time Series Analysis*, 18, 231–51.

——— (1998) Wavelet Shrinkage Denoising using the Non-Negative Garrote. *Journal of Computational and Graphical Statistics*, 7, 469–88.

Gao, W. and Li, B. L. (1993) Wavelet Analysis of Coherent Structures at the Atmosphere-Forest Interface. *Journal of Applied Meteorology*, 32, 1717–25.

Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (1995) *Bayesian Data Analysis*. New York: Chapman & Hall.

Geweke, J. and Porter-Hudak, S. (1983) The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*, 4, 221–38.

Gneiting, T. (2000) Power-Law Correlations, Related Models for Long-Range Dependence, and Their Simulation. *Journal of Applied Probability*, 37, 1104–9.

Godfrey, R. and Rocca, F. (1981) Zero Memory Non-Linear Deconvolution. *Geophysical Prospecting*, 29, 189–228.

Gollmer, S. M., Harshvardhan, Cahalan, R. F. and Snider, J. B. (1995) Windowed and Wavelet Analysis of Marine Stratocumulus Cloud Inhomogeneity. *Journal of the Atmospheric Sciences*, 52, 3013–30.

González-Esparza, J. A., Balogh, A., Forsyth, R. J., Neugebauer, M., Smith, E. J. and Phillips, J. L. (1996) Interplanetary Shock Waves and Large-Scale Structures: Ulysses' Observations in and out of the Ecliptic Plane. *Journal of Geophysical Research*, 101, 17,057–71.

Goupillaud, P., Grossmann, A. and Morlet, J. (1984) Cycle-Octave and Related Transforms in Seismic Signal Analysis. *Geoexploration*, 23, 85–102.

Granger, C. W. J. and Hatanaka, M. (1964) *Spectral Analysis of Economic Time Series*. Princeton: Princeton University Press.

Granger, C. W. J. and Joyeux, R. (1980) An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, 1, 15–29.

Graybill, F. A. (1983) *Matrices with Applications in Statistics* (Second Edition). Belmont, California: Wadsworth.

Green, P. J. and Silverman, B. W. (1994) *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. London: Chapman & Hall.

Greenhall, C. A. (1991) Recipes for Degrees of Freedom of Frequency Stability Estimators. *IEEE Transactions on Instrumentation and Measurement*, 40, 994–9.

Greenhall, C. A., Howe, D. A. and Percival, D. B. (1999) Total Variance, an Estimator of Long-Term Frequency Stability. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 46, 1183–91.

Grossmann, A. and Morlet, J. (1984) Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape. *SIAM Journal on Mathematical Analysis*, 15, 723–36.

Haar, A. (1910) Zur Theorie der Orthogonalen Funktionensysteme. *Mathematische Annalen*, 69, 331–71.

Hamming, R. W. (1989) *Digital Filters* (Third Edition). Englewood Cliffs, New Jersey: Prentice Hall.

Hannan, E. J. (1970) *Multiple Time Series*. New York: John Wiley & Sons.

Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998) *Wavelets, Approximation, and Statistical Applications* (Lecture Notes in Statistics, Volume 129). New York: Springer.

Hess–Nielsen, N. and Wickerhauser, M. V. (1996) Wavelets and Time-Frequency Analysis. *Proceedings of the IEEE*, 84, 523–40.

Holden, A. V. (1976) *Models of the Stochastic Activity of Neurones* (Lecture Notes in Biomathematics, Volume 12). Berlin: Springer–Verlag.

Hosking, J. R. M. (1981) Fractional Differencing. *Biometrika*, 68, 165–76.

——— (1984) Modeling Persistence in Hydrological Time Series Using Fractional Differencing. *Water Resources Research*, 20, 1898–908.

Hsu, D. A. (1977) Tests for Variance Shift at an Unknown Time Point. *Applied Statistics*, 26, 279–84.

Hudgins, L. H. (1992) Wavelet Analysis of Atmospheric Turbulence. Ph.D. dissertation, University of California, Irvine.

Hudgins, L. H., Friehe, C. A. and Mayer, M. E. (1993) Wavelet Transforms and Atmospheric Turbulence. *Physical Review Letters*, 71, 3279–82.

Hurst, H. E. (1951) Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 770–99.

Hurvich, C. M., Deo, R. and Brodsky, J. (1998) The Mean Squared Error of Geweke and Porter–Hudak's Estimator of the Memory Parameter of a Long-Memory Time Series. *Journal of Time Series Analysis*, 19, 19–46.

Hurvich, C. M. and Tsai, C.-L. (1998) A Crossvalidatory AIC for Hard Wavelet Thresholding in Spatially Adaptive Function Estimation. *Biometrika*, 85, 701–10.

Inclán, C. and Tiao, G. C. (1994) Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance. *Journal of the American Statistical Association*, 427, 913–23.

Jansen, M., Malfait, M. and Bultheel, A. (1997) Generalized Cross Validation for Wavelet Thresholding. *Signal Processing*, 56, 33–44.

Jawerth, B. and Sweldens, W. (1994) An Overview of Wavelet Based Multiresolution Analyses. *SIAM Review*, 36, 377–412.

Jenkins, G. M. and Watts, D. G. (1968) *Spectral Analysis and Its Applications*. San Francisco: Holden–Day.

Jensen, M. J. (1999a) Using Wavelets to Obtain a Consistent Ordinary Least Squares Estimator of the Long-Memory Parameter. *Journal of Forecasting*, 18, 17–32.

——— (1999b) An Approximate Wavelet MLE of Short- and Long-Memory Parameters. *Studies in Nonlinear Dynamics and Econometrics*, 3, 239–53.

——— (2000) An Alternative Maximum Likelihood Estimator of Long-Memory Processes Using Compactly Supported Wavelets. *Journal of Economic Dynamics and Control*, 24, 361–87.

Jiang, J., Zhang, D. and Fraedrich, K. (1997) Historic Climate Variability of Wetness in East China (960–1992): A Wavelet Analysis. *International Journal of Climatology*, 17, 969–81.

Johnson, G. E. (1994) Constructions of Particular Random Processes. *Proceedings of the IEEE*, 82, 270–85.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994) *Continuous Univariate Distributions, Volume 1* (Second Edition). New York: John Wiley & Sons.

Johnstone, I. M. and Silverman, B. W. (1997) Wavelet Threshold Estimators for Data with Correlated Noise. *Journal of the Royal Statistical Society, Series B*, 59, 319–51.

Kaplan, L. M. and Kuo, C.-C. J. (1993) Fractal Estimation from Noisy Data via Discrete Fractional Gaussian Noise (DFGN) and the Haar Basis. *IEEE Transactions on Signal Processing*, 41, 3554–62.

——— (1994) Extending Self-Similarity for Fractional Brownian Motion. *IEEE Transactions on Signal Processing*, 42, 3526–30.

Kashyap, R. L. and Eom, K.-B. (1988) Estimation in Long-Memory Time Series Model. *Journal of Time Series Analysis*, 9, 35–41.

Kay, S. M. (1981) Efficient Generation of Colored Noise. *Proceedings of the IEEE*, 69, 480–1.

——— (1988) *Modern Spectral Estimation: Theory & Application*. Englewood Cliffs, New Jersey: Prentice Hall.

Komm, R. W., Gu, Y., Hill, F., Stark, P. B. and Fodor, I. K. (1999) Multitaper Spectral Analysis and Wavelet Denoising Applied to Helioseismic Data. *Astrophysical Journal*, 519, 407–21.

Koopmans, L. H. (1974) *The Spectral Analysis of Time Series*. New York: Academic Press.

Kumar, P. and Foufoula–Georgiou, E. (1993) A Multicomponent Decomposition of Spatial Random Fields 1: Segregation of Large- and Small-Scale Features Using Wavelet Transforms. *Water Resources Research*, 29, 2515–32.

——— (1994) Wavelet Analysis in Geophysics: An Introduction. In *Wavelets in Geophysics*, edited by E. Foufoula–Georgiou and P. Kumar. San Diego: Academic Press, 1–43.

——— (1997) Wavelet Analysis for Geophysical Applications. *Reviews of Geophysics*, 35, 385–412.

Kuo, C., Lindberg, C. and Thomson, D. J. (1990) Coherence Established Between Atmospheric Carbon Dioxide and Global Temperature. *Nature*, 343, 709–14.

Lai, M.-J. (1995) On the Digital Filter Associated with Daubechies' Wavelets. *IEEE Transactions on Signal Processing*, 43, 2203–5.

Lang, M., Guo, H., Odegard, J. E., Burrus, C. S. and Wells, R. O. (1995) Nonlinear Processing of a Shift Invariant DWT for Noise Reduction. In *Wavelet Applications II* (Proceedings of the SPIE 2491), edited by H. H. Szu. Bellingham, Washington: SPIE Press, 640–51.

——— (1996) Noise Reduction using an Undecimated Discrete Wavelet Transform. *IEEE Signal Processing Letters*, 3, 10–12.

Lanning, E. N. and Johnson, D. M. (1983) Automated Identification of Rock Boundaries: An Application of the Walsh Transform to Geophysical Well-Log Analysis. *Geophysics*, 48, 197–205.

Lee, P. M. (1989) *Bayesian Statistics: An Introduction*. London: Edward Arnold.

Liang, J. and Parks, T. W. (1996) A Translation-Invariant Wavelet Representation Algorithm with Applications. *IEEE Transactions on Signal Processing*, 44, 225–32.

Lindsay, R. W., Percival, D. B. and Rothrock, D. A. (1996) The Discrete Wavelet Transform and the Scale Analysis of the Surface Properties of Sea Ice. *IEEE Transactions on Geoscience and Remote Sensing*, 34, 771–87.

Liu, P. C. (1994) Wavelet Spectrum Analysis and Ocean Wind Waves. In *Wavelets in Geophysics*, edited by E. Foufoula–Georgiou and P. Kumar. San Diego: Academic Press, 151–66.

Lugannani, R. and Rice, S. (1980) Saddle Point Approximation for the Distribution of the Sum of Independent Random Variables. *Advances in Applied Probability*, 12, 475–90.

Mak, M. (1995) Orthogonal Wavelet Analysis: Interannual Variability in the Sea Surface Temperature. *Bulletin of the American Meteorological Society*, 76, 2179–86.

Mallat, S. G. (1989a) Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(R)$. *Transactions of the American Mathematical Society*, 315, 69–87.

——— (1989b) A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 674–93.

——— (1989c) Multifrequency Channel Decompositions of Images and Wavelet Models. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37, 2091–110.

——— (1998) *A Wavelet Tour of Signal Processing*. San Diego: Academic Press.

Mallat, S. G. and Zhang, Z. (1993) Matching Pursuits with Time-Frequency Dictionaries. *IEEE Transactions on Signal Processing*, 41, 3397–415.

Mandelbrot, B. B. and van Ness, J. W. (1968) Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review*, 10, 422–37.

Mandelbrot, B. B. and Wallis, J. R. (1968) Noah, Joseph, and Operational Hydrology. *Water Resources Research*, 4, 909–18.

Marple, S. L., Jr. (1987) *Digital Spectral Analysis with Applications*. Englewood Cliffs, New Jersey: Prentice-Hall.

Masry, E. (1993) The Wavelet Transform of Stochastic Processes with Stationary Increments and its Application to Fractional Brownian Motion. *IEEE Transactions on Information Theory*, 39, 260–4.

McCoy, E. J. and Walden, A. T. (1996) Wavelet Analysis and Synthesis of Stationary Long-Memory Processes. *Journal of Computational and Graphical Statistics*, 5, 26–56.

McCoy, E. J., Walden, A.T. and Percival, D. B. (1998) Multitaper Spectral Estimation of Power Law Processes. *IEEE Transactions on Signal Processing*, 46, 655–68.

McHardy, I. and Czerny, B. (1987) Fractal X-Ray Time Variability and Spectral Invariance of the Seyfert Galaxy NGC5506. *Nature*, 325, 696–8.

McLeod, A. I. and Hipel, K. W. (1978) Simulation Procedures for Box–Jenkins Models. *Water Resources Research*, 14, 969–75.

Medley, M., Saulnier, G. and Das, P. (1994) Applications of the Wavelet Transform in Spread Spectrum Communications Systems. In *Wavelet Applications* (Proceedings of the SPIE 2242), edited by H. H. Szu. Bellingham, Washington: SPIE Press, 54–68.

Meyers, S. D., Kelly, B. G. and O'Brien, J. J. (1993) An Introduction to Wavelet Analysis in Oceanography and Meteorology: With Applications to the Dispersion of Yanai Waves. *Monthly Weather Review*, 121, 2858–66.

Mood, A. M., Graybill, F. A. and Boes, D. C. (1974) *Introduction to the Theory of Statistics* (Third Edition). New York: McGraw–Hill.

Morettin, P. A. (1981) Walsh Spectral Analysis. *SIAM Review*, 23, 279–91.

Moulin, P. (1994) Wavelet Thresholding Techniques for Power Spectrum Estimation. *IEEE Transactions on Signal Processing*, 42, 3126–36.

——— (1995) Model Selection Criteria and the Orthogonal Series Method for Function Estimation. In *Proceedings of 1995 IEEE International Symposium on Information Theory*. New York: IEEE Press, 252.

Mulcahy, C. (1996) Plotting and Scheming with Wavelets. *Mathematics Magazine*, 69, 323–43.

Munk, W. H. and MacDonald, G. J. F. (1975) *The Rotation of the Earth*. Cambridge, UK: Cambridge University Press.

Musha, T. and Higuchi, H. (1976) The $1/f$ Fluctuation of a Traffic Current on an Expressway. *Japanese Journal of Applied Physics*, 15, 1271–5.

Nason, G. P. (1996) Wavelet Shrinkage using Cross-Validation. *Journal of the Royal Statistical Society, Series B*, 58, 463–79.

Nason, G. P. and Silverman, B. W. (1995) The Stationary Wavelet Transform and Some Statistical Applications. In *Wavelets and Statistics* (Lecture Notes in Statistics, Volume 103), edited by A. Antoniadis and G. Oppenheim. New York: Springer–Verlag, 281–99.

Ogden, R. T. (1997) *Essential Wavelets for Statistical Applications and Data Analysis*. Boston: Birkhäuser.

Oppenheim, A. V. and Schafer, R. W. (1989) *Discrete-Time Signal Processing*. Englewood Cliffs, New Jersey: Prentice Hall.

Papoulis, A. (1991) *Probability, Random Variables, and Stochastic Processes* (Third Edition). New York: McGraw–Hill.

Percival, D. B. (1983) The Statistics of Long Memory Processes. Ph.D. dissertation, Department of Statistics, University of Washington, Seattle.

——— (1991) Characterization of Frequency Stability: Frequency-Domain Estimation of Stability Measures. *Proceedings of the IEEE*, 79, 961–72.

——— (1992) Simulating Gaussian Random Processes with Specified Spectra. *Computing Science and Statistics*, 24, 534–8.

——— (1995) On Estimation of the Wavelet Variance. *Biometrika*, 82, 619–31.

Percival, D. B. and Constantine, W. (2000) Wavelet-Based Simulation of Stochastic Processes. Technical Report, University of Washington, Seattle.

Percival, D. B. and Guttorp, P. (1994) Long-Memory Processes, the Allan Variance and Wavelets. In *Wavelets in Geophysics*, edited by E. Foufoula–Georgiou and P. Kumar. San Diego: Academic Press, 325–44.

Percival, D. B. and Mofjeld, H. (1997) Analysis of Subtidal Coastal Sea Level Fluctuations Using Wavelets. *Journal of the American Statistical Association*, 92, 868–80.

Percival, D. B., Raymond, G. M. and Bassingthwaighte, J. B. (2000a) Approximation of Stationary Gaussian Processes via Spectral Synthesis Methods. Technical Report, University of Washington, Seattle.

Percival, D. B., Sardy, S. and Davison, A. C. (2000b) Wavestrapping Time Series: Adaptive Wavelet-Based Bootstrapping. In *Nonlinear and Nonstationary Signal Processing*, edited by W. J. Fitzgerald, R. L. Smith, A. T. Walden and P. C. Young. Cambridge, UK: Cambridge University Press, 442–71.

Percival, D. B. and Walden, A. T. (1993) *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*. Cambridge, UK: Cambridge University Press.

Pesquet, J.-C., Krim, H. and Carfantan, H. (1996) Time-Invariant Orthonormal Wavelet Representations. *IEEE Transactions on Signal Processing*, 44, 1964–70.

Peterson, S. (1996) Filtering and Wavelet Regression Methods with Application to Exercise ECG. Ph.D. dissertation, Department of Mathematical Statistics, Lund University, Lund, Sweden.

Popper, W. (1951) *The Cairo Nilometer*, Volume 12 of *University of California Publications in Semitic Philology*. Berkeley, California: University of California Press.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992) *Numerical Recipes in Fortran* (Second Edition). Cambridge, UK: Cambridge University Press.

Priestley, M. B. (1981) *Spectral Analysis and Time Series*. London: Academic Press.

Rabiner, L. R. (1989) A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77, 257–86.

Rao, C. R. and Mitra, S. K. (1971) *Generalized Inverse of Matrices and Its Applications*. New York: John Wiley & Sons.

Rao, K. R. and Yip, P. (1990) *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Boston: Academic Press.

Reid, N. (1988) Saddlepoint Methods and Statistical Inference. *Statistical Science*, 3, 213–38.

Riedel, K. S. and Sidorenko, A. (1995) Minimum Bias Multiple Taper Spectral Estimation. *IEEE Transactions on Signal Processing*, 43, 188–95.

Rioul, O. (1992) Simple Regularity Criteria for Subdivision Schemes. *SIAM Journal on Mathematical Analysis*, 23, 1544–76.

Rioul, O. and Duhamel, P. (1992) Fast Algorithms for Discrete and Continuous Wavelet Transforms. *IEEE Transactions on Information Theory*, 38, 569–86.

Rioul, O. and Vetterli, M. (1991) Wavelets and Signal Processing. *IEEE Signal Processing Magazine*, 8, 14–38.

Robinson, E. A. (1980) *Physical Applications of Stationary Time-Series*. New York: MacMillan.

Robinson, E. A. and Treitel, S. (1980) *Geophysical Signal Analysis*. Englewood Cliffs, New Jersey: Prentice Hall.

Robinson, P. M. (1995) Log-Periodogram Regression of Time Series with Long Range Dependence. *Annals of Statistics*, 23, 1048–72.

Rutman, J. (1978) Characterization of Phase and Frequency Instabilities in Precision Frequency Sources: Fifteen Years of Progress. *Proceedings of the IEEE*, 66, 1048–75.

Scargle, J. D. (1997) Wavelet Methods in Astronomical Time Series Analysis. In *Applications of Time Series Analysis in Astronomy and Meteorology*, edited by T. Subba Rao, M. B. Priestley and O. Lessi. London: Chapman & Hall, 226–48.

Schick, K. L. and Verveen, A. A. (1974) $1/f$ Noise with a Low Frequency White Noise Limit. *Nature*, 251, 599–601.

Serroukh, A. and Walden, A. T. (2000a) Wavelet Scale Analysis of Bivariate Time Series I: Motivation and Estimation. *Journal of Nonparametric Statistics*, 13, 1-36.

——— (2000b) Wavelet Scale Analysis of Bivariate Time Series II: Statistical Properties for Linear Processes. *Journal of Nonparametric Statistics*, 13, 37–56.

Serroukh, A., Walden, A. T. and Percival, D. B. (2000) Statistical Properties and Uses of the Wavelet Variance Estimator for the Scale Analysis of Time Series. *Journal of the American Statistical Association*, 95, 184–96.

Shensa, M. J. (1992) The Discrete Wavelet Transform: Wedding the à Trous and Mallat Algorithms. *IEEE Transactions on Signal Processing*, 40, 2464–82.

Shimomai, T., Yamanaka, M. D. and Fukao, S. (1996) Application of Wavelet Analysis to Wind Disturbances Observed with MST Radar Techniques. *Journal of Atmospheric and Terrestrial Physics*, 58, 683–96.

Sinai, Y. G. (1976) Self-Similar Probability Distributions. *Theory of Probability and Its Applications*, 21, 64–80.

Stein, C. M. (1981) Estimation of the Mean of a Multivariate Normal Distribution. *Annals of Statistics*, 9, 1135–51.

Stoffer, D. S. (1991) Walsh–Fourier Analysis and Its Statistical Applications. *Journal of the American Statistical Association*, 86, 461–85.

Stoffer, D. S., Scher, M. S., Richardson, G. A., Day, N. L. and Coble, P. A. (1988) A Walsh–Fourier Analysis of the Effects of Moderate Maternal Alcohol Consumption on Neonatal Sleep-State Cycling. *Journal of the American Statistical Association*, 83, 954–63.

Stollnitz, E. J., DeRose, T. D. and Salesin, D. H. (1996) *Wavelets for Computer Graphics: Theory and Applications*. San Francisco: Morgan Kaufmann.

Strang, G. (1989) Wavelets and Dilation Equations: A Brief Introduction. *SIAM Review*, 31, 614–27.

——— (1993) Wavelet Transforms Versus Fourier Transforms. *Bulletin of the American Mathematical Society*, 28, 288–305.

Strang, G. and Nguyen, T. (1996) *Wavelets and Filter Banks*. Wellesley, Massachusetts: Wellesley–Cambridge Press.

Strichartz, R. S. (1993) How to Make Wavelets. *American Mathematical Monthly*, 100, 539–56.

——— (1994) Construction of Orthonormal Wavelets. In *Wavelets: Mathematics and Applications*, edited by J. J. Benedetto and M. W. Frazier. Boca Raton, Florida: CRC Press, 23–50.

Sweldens, W. (1996) The Lifting Scheme: A Custom-Design Construction of Biorthogonal Wavelets. *Applied and Computational Harmonic Analysis*, 3, 186–200.

Taswell, C. (1996) Satisficing Search Algorithms for Selecting Near-Best Bases in Adaptive Tree-Structured Wavelet Transforms. *IEEE Transactions on Signal Processing*, 44, 2423–38.

——— (2000) Constraint-Selected and Search-Optimized Families of Daubechies Wavelet Filters Computable by Spectral Factorization. *Journal of Computational and Applied Mathematics*, 121, 179–95.

Taswell, C. and McGill, K. (1994) Algorithm 735: Wavelet Transform Algorithms for Finite-Duration Discrete-Time Signals. *ACM Transactions on Mathematical Software*, 20, 398–412.

Teichroew, D. (1957) The Mixture of Normal Distributions with Different Variances. *Annals of Mathematical Statistics*, 28, 510–2.

Tewfik, A. H. and Kim, M. (1992) Correlation Structure of the Discrete Wavelet Coefficients of Fractional Brownian Motion. *IEEE Transactions on Information Theory*, 38, 904–9.

Tewfik, A. H., Kim, M. and Deriche, M. (1993) Multiscale Signal Processing Techniques: A Review. In *Signal Processing and its Applications*, Volume 10 of *Handbook of Statistics*, edited by N. K. Bose and C. R. Rao. Amsterdam: North–Holland, 819–81.

Thomson, D. J. (1982) Spectrum Estimation and Harmonic Analysis. *Proceedings of the IEEE*, 70, 1055–96.

Torrence, C. and Compo, G. P. (1998) A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79, 61–78.

Toussoun, O. (1925) Mémoire sur l'Histoire du Nil. In *Mémoires a l'Institut d'Égypte*, 9, 366–404.

Tukey, J. W. (1977) *Exploratory Data Analysis*. Reading, Massachusetts: Addison–Wesley.

Unser, M. (1995) Texture Classification and Segmentation Using Wavelet Frames. *IEEE Transactions on Image Processing*, 4, 1549–60.

Van Trees, H. L. (1968) *Detection, Estimation, and Modulation Theory*, Part I. New York: John Wiley & Sons.

Vetterli, M. and Herley, C. (1992) Wavelets and Filter Banks: Theory and Design. *IEEE Transactions on Signal Processing*, 40, 2207–32.

Vidakovic, B. (1998) Nonlinear Wavelet Shrinkage with Bayes Rules and Bayes Factors. *Journal of the American Statistical Association*, 93, 173–9.

——— (1999) *Statistical Modeling by Wavelets*. New York: John Wiley & Sons.

Wahba, G. (1980) Automatic Smoothing of the Log Periodogram. *Journal of American Statistical Association*, 75, 122–32.

Walden, A. T. (1985) Non-Gaussian Reflectivity, Entropy, and Deconvolution. *Geophysics*, 50, 2862–88.

——— (1990) Variance and Degrees of Freedom of a Spectral Estimator Following Data Tapering and Spectral Smoothing. *Signal Processing*, 20, 67–79.

——— (1994) Interpretation of Geophysical Borehole Data via Interpolation of Fractionally Differenced White Noise. *Applied Statistics*, 43, 335–45.

Walden, A. T. and Contreras Cristan, A. (1998a) The Phase-Corrected Undecimated Discrete Wavelet Packet Transform and its Application to Interpreting the Timing of Events. *Proceedings of the Royal Society of London, Series A*, 454, 2243–66.

——— (1998b) Matching Pursuit by Undecimated Discrete Wavelet Transform for Non-Stationary Time Series of Arbitrary Length. *Statistics and Computing*, 8, 205–19.

Walden, A. T., McCoy, E. J. and Percival, D. B. (1995) The Effective Bandwidth of a Multitaper Spectral Estimator. *Biometrika*, 82, 201–14.

Walden, A. T., Percival, D. B. and McCoy, E. J. (1998) Spectrum Estimation by Wavelet Thresholding of Multitaper Estimators. *IEEE Transactions on Signal Processing*, 46, 3153–65.

Wang, Y. (1996) Function Estimation via Wavelet Shrinkage for Long-Memory Data. *Annals of Statistics*, 24, 466–84.

Welch, P. D. (1967) The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15, 70–3.

Whitcher, B. J (1998) Assessing Nonstationary Time Series Using Wavelets. Ph.D. dissertation, Department of Statistics, University of Washington, Seattle.

Whitcher, B. J, Byers, S. D., Guttorp, P. and Percival, D. B. (2000a) Testing for Homogeneity of Variance in Time Series: Long Memory, Wavelets and the Nile River. Technical Report, National Research Center for Statistics and the Environment, University of Washington, Seattle.

Whitcher, B. J, Guttorp, P. and Percival, D. B. (2000b) Wavelet Analysis of Covariance with Application to Atmospheric Time Series. *Journal of Geophysical Research*, 105, 14,941–62.

Wickerhauser, M. V. (1994) *Adapted Wavelet Analysis from Theory to Software*. Wellesley, Massachusetts: A K Peters.

Wood, A. T. A. and Chan, G. (1994) Simulation of Stationary Gaussian Processes in $[0,1]^d$. *Journal of Computational and Graphical Statistics*, 3, 409–32.

Wornell, G. W. (1990) A Karhunen–Loève-like Expansion for $1/f$ Processes via Wavelets. *IEEE Transactions on Information Theory*, 36, 859–61.

——— (1993) Wavelet-Based Representations for the $1/f$ Family of Fractal Processes. *Proceedings of the IEEE*, 81, 1428–50.

——— (1996) *Signal Processing with Fractals: A Wavelet-Based Approach*. Upper Saddle River, New Jersey: Prentice Hall.

Wornell, G. W. and Oppenheim, A. V. (1992) Estimation of Fractal Signals from Noisy Measurements using Wavelets. *IEEE Transactions on Signal Processing*, 40, 611–23.

Yaglom, A. M. (1958) Correlation Theory of Processes with Random Stationary *n*th Increments. *American Mathematical Society Translations* (Series 2), 8, 87–141.

Zurbuchen, Th., Bochsler, P. and von Steiger, R. (1996) Coronal Hole Differential Rotation Rate Observed With SWICS/Ulysses. In *Solar Wind Eight*, edited by D. Winterhalter, J. T. Gosling, S. R. Habbal, W. S. Kurth and M. Neugebauer. Woodbury, New York: American Institute of Physics, 273–6.



[General Information]
□□=□□□□□□□□□□□
□□□=□□□□□□
□□=□□□DONALD B.PERCIVAL □□□ANDREW T.WALDEN
□ □□□□□
□□=562
SS□=11532376
□□□□=2006□03□□1□
□□□=□□□□□□□
□□=24cm
□□□□=69.00
□□□=□□□□(□□: □□ □□: □□□□□□)□□□□ □□□□□□
□□□□□□=□□□Donald B.Percival□□□□Andrew T.Walden□□□□□□□.□□□□□□□□□□□□.□□□□□□□,2006.03.
□□□□=□□11□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□